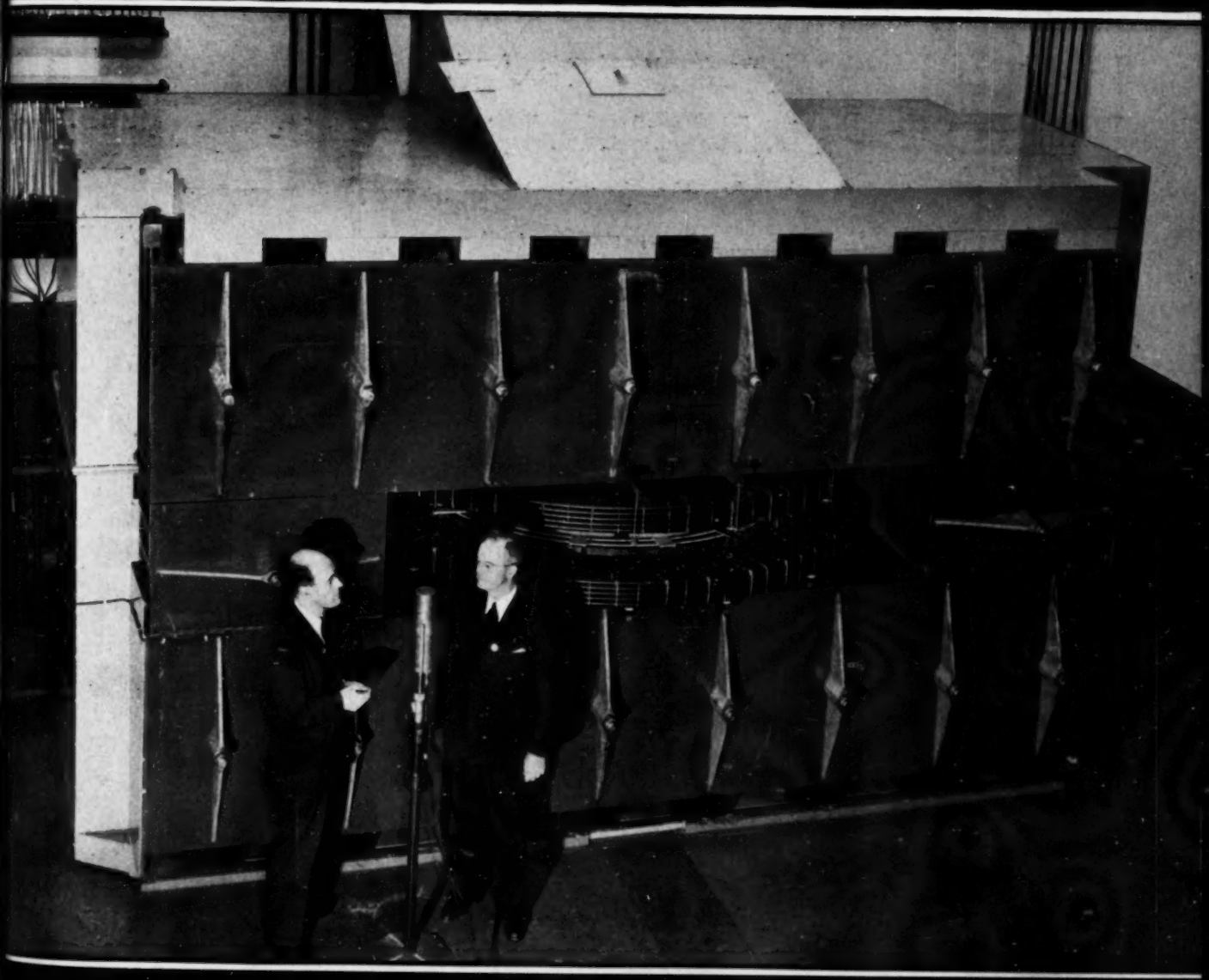
JOURNAL OF APPLIED PHYSICS



JOURNAL OF APPLIED PHYSICS

Vol. 16, No. 10 OCTOBER, 1945

In This Issue

Page

Basic Principles of Semi-Conductors FREDERICK SEITZ 553

Electrical Properties of Semi-Conductors ROBERT J. MAURER 563

On the Representation of Rigid Rotations H. SCHWERDTFEGER 571

Contributed Original Research

100-Million Volt Induction Electron Accelerator W. F. WESTENDORP AND E. E. CHARLTON 581

Thermal Expansion and Second-Order Transition Effects in High Polymers R. F. BOYER AND R. S. SPENCER 594

Method of Measuring the Radiofrequency Resistance of Wires CHANDLER STEWART, JR. 608

Absorption Displacement in X-Ray Diffraction by Cylindrical Samples B. E. WARREN 614

Molecular Weight Studies on High Polymers with the Electron Microscope R. F. BOYER AND R. D. HEIDENREICH 621

Conversion Diagrams for Triode Tube Mixers HARRY STOCKMAN 639

Letters to the Editor:

On a Novel Form of Refrigerator W. D. HAYES 642

Wave Equations for Finite Elastic Strains HUBERT M. JAMES AND EUGENE GUTH 643

In Every Issue

Here and There 577

New Booklets 580

Recent Applications of Physics 644

Index to Advertisers xxxvi

BOARD OF EDITORS

ELMER HUTCHISSON, *Editor*

Associate Editors

R. B. BARNES
OTTO BEECK
W. F. BUSSE
S. H. CALDWELL
CLARK GOODMAN
SELIG HECHT
W. J. LYONS
MORRIS MUSKAT
THOMAS H. OSGOOD
PAUL E. SABINE
FREDERICK SEITZ
J. R. TOWNSEND

Publication of the American Institute of Physics

ADMINISTRATIVE STAFF

HENRY A. BARTON, *Director*
MARGARET S. GRIFFIN, *Publications Manager*
THEODORE VORBURGER, *Business Manager*
JOHN T. TATE, *Adviser on Publications*

GOVERNING BOARD

PAUL E. KLOPSTEG, *Chairman*
WALLACE WATERFALL, *Secretary*
GEORGE B. PEGRAM, *Treasurer*
T. D. COPE
R. B. DOW
L. A. DUBRIDGE
FLOYD A. FIRESTONE
HARVEY FLETCHER
R. C. GIBBS
K. S. GIBSON
HERMANN MARK
A. G. WORTHING
JOHN T. TATE
L. W. TAYLOR
N. W. TAYLOR
WALLACE WATERFALL

The JOURNAL OF APPLIED PHYSICS, published monthly at Prince and Lemon Sts., Lancaster, Pa., is devoted to physics in its role as the science basic to other natural sciences and to the arts and industries. Previous to 1937 this publication was known as PHYSICS. The JOURNAL OF APPLIED PHYSICS publishes editorials, news of physicists, and reviews, as well as technical papers of applied physics.

Manuscripts should be submitted to Elmer Hutchisson, Editor, Case School of Applied Science, Cleveland 6, Ohio.

Subscriptions and orders for back numbers should be addressed to Prince and Lemon Streets, Lancaster, Pa. or to the American Institute of Physics, 57 East 55 Street, New York 22, N. Y.

Subscription Price

	U. S. AND CANADA	ELSEWHERE
To members of scientific societies	$5.00	$5.70
To all others	7.00	7.70

Back Number Prices

Complete set: Vol. 1, July, 1931–Vol. 15, 1944—$94.50
Yearly back number rate: $7.70
Single copies: $0.70 each.

Changes of address and proofs should be addressed to the Publications Manager.

Advertising rates supplied on request. Orders, advertising copy, and cuts should be sent to the American Institute of Physics.

Entered as second class matter January 22, 1937, at the Post Office at Lancaster, Pennsylvania, under the Act of March 3, 1879
Accepted for mailing at the special rate of postage provided for in the Act of February 28, 1925, authorized May 2, 1932.

Copyright 1945, by the American Institute of Physics.

Journal
of
Applied Physics

Volume 16, Number 10 October, 1945

The Basic Principles of Semi-Conductors[1]

BY FREDERICK SEITZ

Carnegie Institute of Technology, Pittsburgh, Pennsylvania

A review is given of typical deviations from ordinary valence rules by alloys and inorganic compounds. The effect on the electrical properties of a compound of a deviation from stoichiometric proportions is discussed. The value of the information obtained from electrical measurements is illustrated by a discussion of the properties of several semi-conducting compounds.

1. INTRODUCTION[2]

IT is a commonly known fact that the best insulating materials, such as organic resins and hydrocarbons and the high melting inorganic materials like the silicates and light-metal oxides, obey the rules of combining proportions very accurately. On the other hand, there is a large number of non-metallic solids which do not obey the rules of combining proportions very accurately and which, interestingly enough, are usually not very good insulators. In this category, for example, fall materials such as copper oxide, lead sulfide, iron sulfide, and other heavy-metal oxides, sulfides, and selenides. Careful examination shows that these materials are almost inevitably electronic conductors in the vicinity of room temperature, so that they represent a class of substance intermediate between the ideal metals and ideal insulators. We shall follow convention by calling this class *semi-conductors*. It should be added that several monatomic materials, such as selenium and tellurium, fall in the same category in the sense that they are electronic conductors that have conductivities intermediate between the values usually associated with the pure metals and the ideal insulators. These monatomic materials evidently cannot possess compositions which deviate from the rules of combining proportions; however, their electrical properties usually are greatly affected by impurities, so that we may conclude that their electronic conductivity is intimately connected with their chemical behavior.

In the present paper the relationship between the chemical and electrical properties of semi-conductors will be reviewed with the purpose of showing that the two are inextricably interwoven, and actually are manifestations of one set of principles. In addition, it will be shown that, as in many other fields of chemistry, purely physical measurements make it possible to explore with comparative ease a domain that would otherwise have been opened only with the greatest difficutly.

2. THE LAWS OF COMBINING PROPORTIONS

It is an interesting fact that in a large part of the conventional field of inorganic chemistry the

[1] Based on a paper presented at a symposium on semi-conductors at the Detroit meeting of the American Chemical Society, April, 1943.

[2] More detailed surveys of this topic may be found in the following books: F. Seitz, *The Modern Theory of Solids* (McGraw-Hill Book Company, Inc., New York, 1940); N. F. Mott and R. W. Gurney, *Electronic Processes in Ionic Crystals* (Oxford University Press, New York, 1940); A. H. Wilson, *Semi-Conductors and Metals* (Cambridge University Press, New York, 1939).

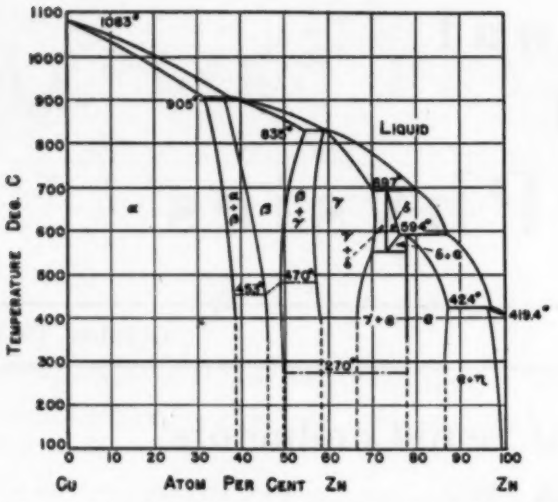


FIG. 1. The phase diagram of the brass system (Cu-Zn). The phases of this typical alloy system form over wide ranges of composition and for combining ratios that have no meaning from the standpoint of conventional valence theory.

amount of each element present in a given compound may be uniquely specified by giving the valence states of each of the elements in the compound. The valences which a given element may exhibit are in turn uniquely limited by its electronic structure or, more practically, by its position in the periodic chart. In this category fall, for example, such simple materials as the common halides and oxides of the lighter metals. The field of organic chemistry, on the other hand, is set apart not only by the fact that the compounds with which it deals are those found in living matter, but also by the fact that the combining ratios are not uniquely determined by stating the valence of the elements contained in these compounds. We know, as a result of the work of Kekule and his followers, that the principles of valency actually do apply in this field; however, it is necessary to augment the principles of valency with a theory of molecular structure before the combining proportions of a compound can be specified. Once the structure pattern of a compound has been settled upon, the laws of valency operate as effectively as in the part of the field of inorganic chemistry referred to above. In fact, once the importance of structure has been realized, we are able to correlate a much larger part of the field of inorganic chemistry with the use of simple principles of valency. For example, a clear understanding of the nature and composition of the many complex crystalline silicates and similar "chain" compounds can be then obtained.

It is a notable fact that the rules of combination obeyed by ionic and homopolar compounds are identical insofar as obedience to the laws of valency are concerned. The principal difference between the two types of compound shows up in the nature of the structure found in each. The great similarity goes back, of course, to the fact that the ability of an atom to become charged and its ability to form homopolar bonds are intimately related to the structure of its outer shell, which in turn determines its valence.

In spite of the degree to which the simple theory of valency is successful in determining combining proportions in large parts of the field of chemistry, when coupled with the rules of structure, we need not look very far to find an important class of simple crystalline compound in which the theory fails completely. It is evident (Fig. 1) that such alloys as the brasses (Cu-Zn alloys) and the bronzes (Cu-Sn alloys) obey an entirely different set of laws of combining proportions. For, in the first place, many of the metallic elements combine over a wide range of composition; and, in the second, even if we overlook the range of solubility, we must introduce formulae of the type Cu_5Zn_8, which are meaningless from ordinary valence theory, in attempting to describe the composition of the alloys. It is true that liquids and glasses usually show a similar range of solubility; however, these materials are in a somewhat different category since they are not crystalline and since they are formed by dissolving a compound which obeys the ordinary rules of combining proportions in another which does the same.

A significant clue to the irregular position occupied by the alloys is provided by the fact that the metals which alloy well, in the sense that their phases have well-separated solubility limits, are those which lie close to one another in the electromotive series. This means that they have very nearly the same chemical affinity, and as a result do not become strongly charged when in combination with one another. Thus, the ion-ion forces which play a very important role in most inorganic compounds are weak in the alloys. A careful analysis shows that in ionic substances which are composed of elements lying far apart in the electromotive series, these forces actually are responsible for the ordinary rules of combining proportions. For they require that the system be electrostatically neutral and, as a result, require that there be fixed ratios of the electropositive and electro-

negative elements. The same is undoubtedly true to a large extent in many of the homopolar compounds in which the homopolar bonds occur between different elements; that is, the displacement of charge from one atom to another that takes place in such compounds is probably sufficient to compel them to obey the rules followed by ionic substances. In addition, of course, the properties of the homopolar bond are such that the rules of valence are closely satisfied. We may conclude, then, that the constituents of ideal alloys disobey the ordinary rules of valence chemistry because (*a*) they lie so close together in the electromotive series that they do not obtain sufficient ionic charge to behave like ionic crystals and (*b*) the bond between them is sufficiently unlike the homopolar bond that they need not follow the same rules as homopolar substances.

Further insight into the rules obeyed by ionic and homopolar compounds on the one hand and by the alloys on the other may be obtained by consideration of the principles that determine the stability of any chemical system. According to thermodynamics the stable state of any system is determined by the condition that its free energy is a minimum. Since we are primarily interested in solids, we may take the free energy in the form

$$A = E - TS. \tag{1}$$

Here E is the mechanical energy of the system, which is equal to the sum of the kinetic and potential energy of its constituent particles; S is its entropy; and T is the temperature, which for convenience we may take to be the absolute temperature. It is well known that the entropy of any given energy state of a system may be related to the number of possible ways N in which it may have that energy by means of the equation:

$$S = kT \log N, \tag{2}$$

in which k is Boltzmann's constant. Thus S increases as the number of ways in which a given energy state may be achieved increases.

We may now note two extreme cases in which the free energy is used to determine the equilibrium state of a system. In the first place, let us suppose we are at the absolute zero of temperature; then T is zero and A reduces to E. Thus, in this circumstance, the condition for chemical equilibrium is identical with the condition for mechanical equilibrium. On the other hand, let us suppose that the temperature is so high that E in Eq. (1) is negligible in comparison with TS. Then A is a minimum when S is a maximum; that is, the system tends to that energy state which has the greatest value of N associated with it or which has the greatest degree of randomness. At any intermediate temperature in which E and TS are comparable, the equilibrium state will be determined by a compromise between the tendency of E to become as small as possible and S to become as large as possible. This competition is usually very genuine, for the states of low energy usually have small values of N or S associated with them, whereas, the states of large S and N usually have large values of E.

We may now apply these basic principles to the case of solid compounds, restricting our attention at first to the absolute zero of temperature where E alone determines the state of equilibrium. If we form an ionic compound in a case in which we have an excess of one of the constituent elements, experience tells us that the system is most stable, that is, has the lowest energy when "ideal" combining proportions of the two unite to form as much of the salt as possible, and the excess constituent remains uncombined. We know from the theory of ionic crystals that this occurs because the salt has its lowest energy when the ions are very precisely arranged in a perfect lattice. If the excess constituent were forced into the lattice as a solution of neutral atoms or molecules, so much work would have to be done that there would be a resultant rise in energy. Exactly the same situation occurs when we combine the constituents which make up an ideal homopolar compound.

On the other hand, when we form an ideal alloy, the atoms are relatively unconcerned about the precise nature of their neighbors since all of the constituent atoms are close together in the electromotive series and since none exhibit homopolar tendencies. Thus, in this case there is a net decrease in energy when all of the constituents are in the solid, regardless of the composition.

Now we may note that in general as we leave the absolute zero of temperature, the change that takes place in ionic and homopolar solids should push the state of these materials in the direction of that found in the ideal alloys at the absolute zero of temperature. The reason for this is that the TS term in (1), which requires randomness, will begin to exert its effect as soon as T is no longer zero. It is evident that the state of order in the ideal alloy is less than that found

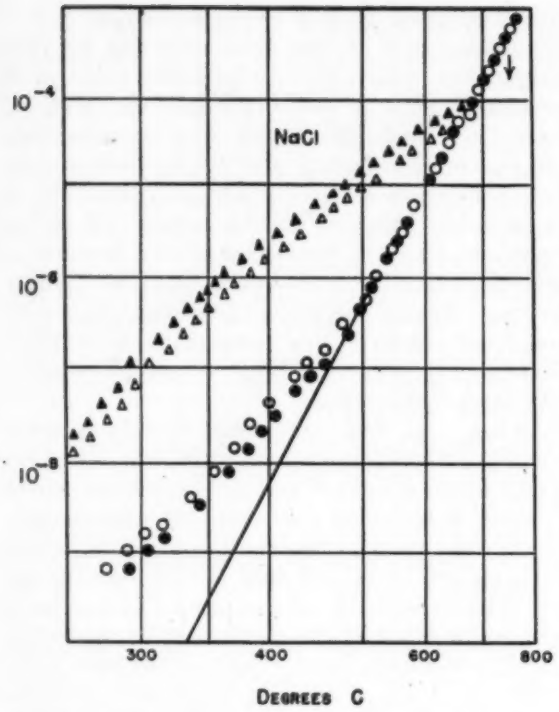


FIG. 2. The electrical conductivity of several specimens of sodium chloride (after Lehfeldt). The vertical scale is in units of ohm^{-1} cm^{-1}. It may be noted that the conductivities of all specimens are the same just below the melting point, but deviate at lower temperatures.

in the well-disciplined arrangements of ionic and homopolar substances. Thus we may begin to expect deviations from ideal stoichiometric relations to take place in these materials at sufficiently high temperatures under proper circumstances, such as if they are prepared with an excess of one of the constituents.

It is interesting to note in passing that under proper conditions even ideal alloys may show an opposite tendency at absolute zero; that is, show a tendency to obey the rules of valency. For example, there is a phase in the brass system (Fig. 1) known as β-brass, which has a body-centered cubic lattice and possesses the approximate composition CuZn. At elevated temperatures this phase forms over a fairly wide range of composition. Moreover, the arrangement of atoms among the various lattice sites of the lattice is random. However, as the temperature is lowered, two changes occur. In the first place, the solubility limits become very small; in the second, the arrangement of atoms becomes *ordered* in the sense that each copper atom prefers to have its eight neighbors be zinc atoms and *vice versa*. Thus, at low temperatures the lattice resembles very closely that of cesium chloride. It has been shown by Mott that this change is a result of the charges which the atoms possess as a result of their relatively small difference in position in the electromotive series; that is, the zinc atom behaves like a positive ion and the copper like a negative ion. The ionic charge is of the order of 0.1 electronic unit.

A more extreme example occurs in the magnesium-antimony system. Both these metals form ideal alloys when combined with metals which are close to them in the electromotive series. However, they do not lie close to one another in this series and as a result form only a single compound having the composition Mg_3Sb_2, which has very narrow solubility limits. It is evident that this compound, though an electronic conductor, has much in common with the ionic substances. In it magnesium behaves like a bivalent metal and antimony like a trivalent electronegative atom.

We shall now examine the types of deviation from stoichiometric proportions that occur in "ideal" ionic and homopolar substances and shall illustrate the way in which these deviations affect the electronic conductivity of these materials.

3. THE ALKALI HALIDES[3]

The alkali halides are the prototypes of the large and important class of ionic crystals and as such merit particular attention. In addition, their stoichiometric properties have been investigated extensively by Pohl and his associates under just the conditions in which we are interested in this article.

Ordinary specimens of the alkali halides which have been prepared under conditions of high chemical purity show no significant deviations from ideal stoichiometric combination. For this reason it might seem at first sight that their lattices are perfect representatives of the ideal ionic structures. If, however, we examine their electrical conductivity, we find good evidence to show that at any finite temperature these salts actually are the seat of an important effect which has bearing on their composition, as was first pointed out by Frenkel. Figure 2 shows the logarithm of the ohmic electrical conductivity of sodium chloride as a function of the reciprocal

[3] The experimental research on the properties of the alkali-halide crystals is described in the books of reference 2 and in the paper by R. W. Pohl, Physik. Zeits. **39**, 36 (1938).

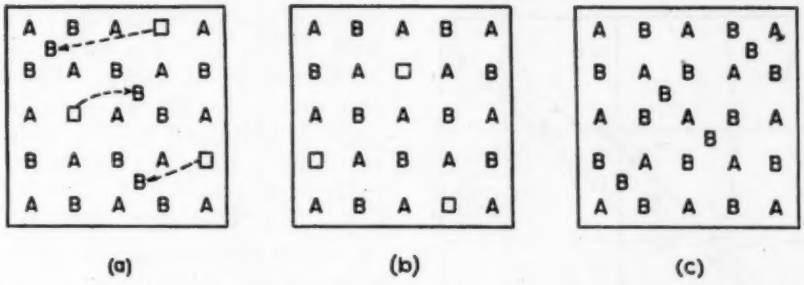


FIG. 3. Three types of lattice defects. (a) represents interstitial atoms and vacancies that are produced by atoms leaving normal positions in the lattice. (b) represents vacancies that are produced as a result of migration of one of the constituents out of the lattice. (c) represents interstitial atoms that are produced as a result of vaporization of one of the constituents from the surface of the lattice and the subsequent diffusion of the extra constituent into the interior of the crystal. It is believed that the normal alkali halides possess equal numbers of vacancies of the type shown in (b). In this case, the atoms that diffuse out of the lattice unite on the surface to form additional layers of the crystal.

of the absolute temperature. Transport experiments show that this conductivity is entirely ionic and is due to the migration of both positive and negative ions in variable ratio. Frenkel emphasized that the occurrence of this ionic conductivity in the solid can only mean that there must be present a degree of disorder in the lattice. For if the lattice were perfect and all ions were rigidly attached to fixed positions, except for thermal oscillations, we would expect no current to flow until the applied electric field reaches a value sufficient to tear the ions from the lattice and drive them to the electrodes. This type of conductivity evidently would not be ohmic, as the observed conductivity is, but would be analogous to dielectric breakdown, in which a large current is generated suddenly when the field reaches a definite value. Frenkel concluded that the ions must be migrating at any temperature at which the ionic conductivity is observed and that the field must play the role of directing this migration. It may readily be shown that such a mechanism will lead to Ohm's law.

It was originally suggested that the ions in the alkali halides occasionally leave their normal lattice positions as the result of thermal fluctuations and force their way into interstitial positions where they are more or less free to wander. (Fig. 3a). It is now known, however, as a result of extensive research, that the mechanism actually is somewhat different. Under the influence of thermal fluctuations, ions in the surface layers jump to the outer surface and leave *vacancies* behind (Figs. 3b, 5a). These vacancies wander through the lattice and permit a degree of mobility to the normal ions in the salt. The higher the temperature, the larger is the number of vacancies and the greater is the conductivity. This situation leads to the type of conductivity shown in Fig. 2 which increases with increasing temperature. It may be seen that the conductivity curve illustrated in the figure is perfectly linear near the melting point. This means that in this range of temperature the conductivity C obeys the equation:

$$C = Ae^{-Q/RT}, \tag{3}$$

in which A and Q are constants that are dependent upon the material, R is the gas constant, and T is the absolute temperature. This equation resembles closely that obeyed by the vapor pressure of a liquid or solid in which case Q is the heat of sublimation. The analogy actually is very close, for in the electrical case C is proportional to the density of vacancies in the lattice; whereas, in the case of vaporization the vapor pressure is proportional to the density of vapor molecules in the gas. It follows that Q in Eq. (3) is related to the energy required to form the vacancies in the crystal, and that these vacancies may be regarded as a type of "vapor" within the solid.

It may be observed that as the temperature is lowered the curve shown in Fig. 2 shows deviations from the linear relation that is valid just below the melting point. The temperature at which this deviation occurs depends upon the previous history of the specimen, although the conductivity at high temperature is not dependent upon this history. It is now known that this deviation is related to the "freezing-in" of vacancies as the temperature is lowered. Impurities may play an important role in aiding

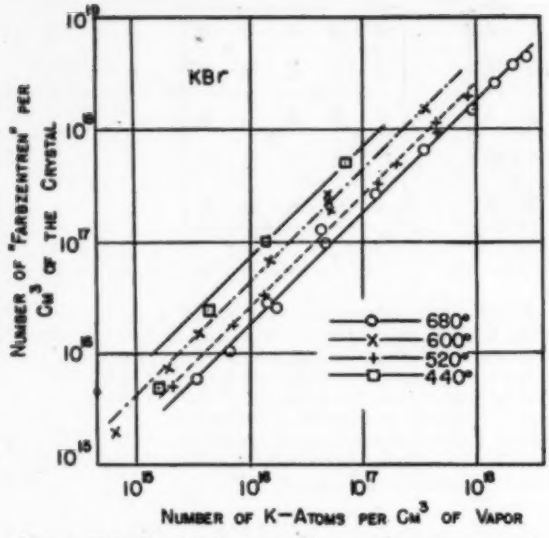


FIG. 4. Relationship between the number of excess potassium atoms dissolved in KBr and the density of atoms in the vapor surrounding the crystal. (After Pohl.)

this freezing, as is shown by the fact that the temperature at which the deviation occurs is lower the higher the purity. In any case, there is good reason to believe that if sufficient time were allowed for true equilibrium to occur at any given temperature, the conductivity of a pure specimen would follow accurately the linear relation observed at high temperatures.

Under normal conditions the numbers of positive-ion vacancies and negative-ion vacancies are equal. Thus, even though the ratio of vacancies to normal sites becomes as large as 0.1 percent near the melting point, no deviation from ideal stoichiometric relations is observed, as would be the case if there were more vacancies of one type than of the other. The reason for this equality in the number of vacancies is clear. If more positive ions than negative ones diffused to the surface, the surface would obtain a resultant positive charge and the interior would obtain a negative charge. This development of electrostatic potential would halt the diffusion process before it had gone far enough to lead to measurable differences in the ratio. It is conceivable that each of the excess halogen ions that would be left in the lattice under these conditions could give up an electron to the alkali metal ions at the surface and thereby prevent the development of potential. In this case the crystals would contain a layer of alkali metal on the surface and an excess of neutral halogen in the interior. Since the layer of alkali metal would be volatile at the temperatures of interest, it would disappear, and the crystals would be left with a stoichiometric excess of halogen. Simple consideration of energy, however, shows that this process could not occur below the melting point in the alkali halides, although, as we shall see later, a process very much like it occurs in other materials.

Suppose now that we heat an alkali-halide crystal in the presence of either the alkali metal vapor or the halogen gas. Then we might expect the gas atoms to take advantage of the porosity of the crystal that results from the presence of the vacancies and diffuse into it. This actually occurs. In other words, it is possible to produce alkali-halide crystals containing a stoichiometric excess of one of the constituents by heating in the presence of that constituent. Crystals containing an excess of the metal have been studied very extensively by Pohl and his co-workers, and a great deal is known about them at present. Figure 4 shows the relationship between the number of potassium atoms in the vapor and the excess in the crystal when KCl is heated in the presence of potassium at several temperatures. It is possible to determine the equilibrium constant for the reaction:

$$K(\text{gas}) \rightleftarrows K(\text{crystal}) \qquad (4)$$

from curves of this type. It turns out that the alkali atom becomes ionized after striking the surface of the crystal and that its valence electron enters the crystal and occupies a halogen-ion vacancy (Fig. 5). Other halogen ions diffuse to the surface to maintain a balance of electrical charge and form layers of the salt. If the temperature is sufficiently high, these electrons may migrate about in the crystal and transform it into an electronic conductor; that is, to a semi-conductor. Thus, we see that in this case there is a

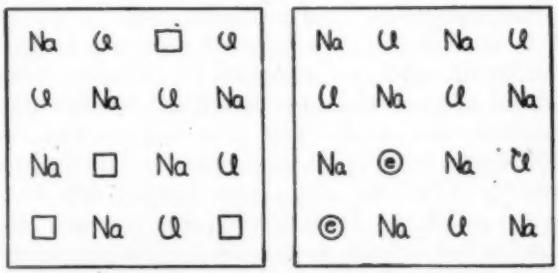


FIG. 5. The figure on the left represents an uncolored sodium chloride crystal in which there are equal numbers of vacancies of each sign. The right-hand figure represents the change produced when excess sodium is added. The positive ion vacancies become filled with sodium ions and the halogen vacancies become filled with electrons.

very intimate relationship between the electronic conductivity and the deviations from ideal proportions that occur when the crystal is heated in the presence of the metal.

The electrons bound at halogen-ion vacancies may absorb optical radiation that lies in the visible part of the spectrum. As a result, they discolor the crystal. When a colored crystal is placed between electrodes at a temperature sufficiently high to promote diffusion, the discoloration diffuses to the anode, and the crystal reverts to its original condition. During this process the electrons diffuse out of the crystal via the anode, and the excess alkali ions are converted to neutral atoms at the cathode. The removal of the stoichiometric excess of metal is accompanied by the decrease in electrical conductivity shown in Fig. 6. The shaded portion corresponds to the electronic part of the current.

It is evident from the observations on the alkali halides that if we were constructing a temperature *versus* composition-phase diagram for this system, it would be improper to represent the salt by an infinitely narrow line. Instead, the true relation should be that shown in Fig. 7 in an exaggerated degree. The boundary limits of the NaCl phase are shown to be fairly symmetrical since it is known that excess halogen can be added to the crystal as well as excess alkali metal.

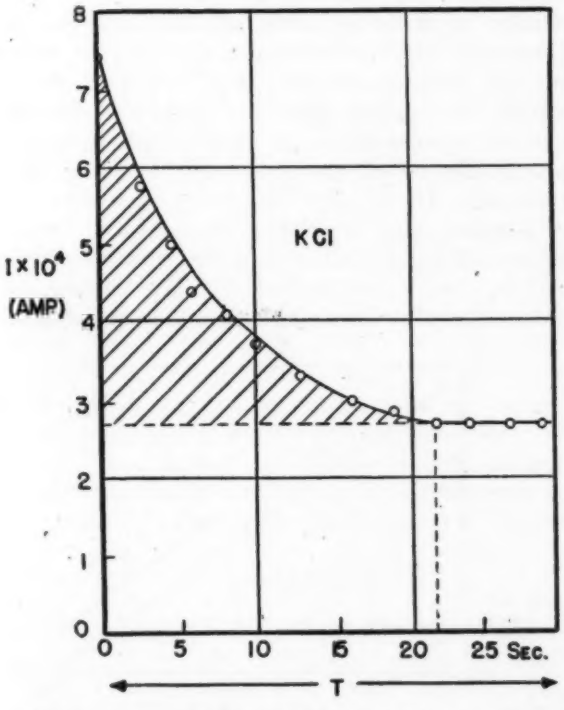


FIG. 6. Decrease in conductivity of colored potassium chloride as the color centers are removed from the lattice by electrolysis. (After Pohl.)

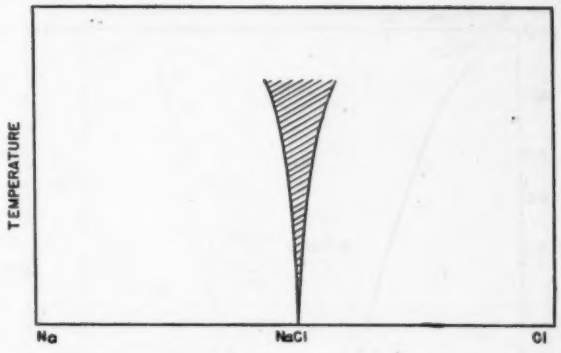


FIG. 7. Schematic phase diagram of the Na-Cl system. The cross-hatched region indicates the range of composition over which the crystal may be formed. The width of this region actually is about 0.1 percent near the melting point instead of the large value shown.

4. THE SILVER HALIDES

Silver chloride and silver bromide resemble the alkali halides closely in that they possess the sodium-chloride structure and are composed of singly charged ions. Moreover, at room temperature they possess a relatively large ionic conductivity which is due almost entirely to the mobility of silver ions. Careful analysis shows, however, that the lattice defects which occur in these salts probably are of a different nature from those found in the alkali halides. It seems most likely that in these cases a fraction of the silver ions force their way into interstitial positions, so that the ionic conductivity is due to the migration, both of these ions and of the vacancies left behind. As a result of this difference, it does not appear to be possible to convert the silver halides into electronic conductors by the methods used for the alkali halides.

It should be added that the high ionic mobility of the silver halides observed at room temperature plays an important role in the processes that take place in the darkening of the photographic emulsion.

5. ZINC OXIDE[4]

Zinc oxide that has been prepared at room temperature by a wet method is a white powder which has negligible conductivity. If, however,

[4] See reference 2 and the many papers by C. Wagner and his associates; for example, Zeits f. physik. Chemie **B22**, 199 (1933).

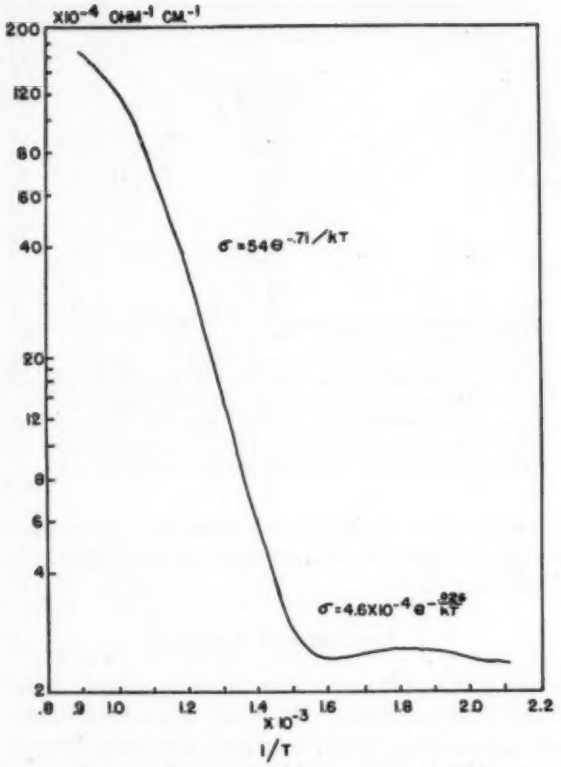


FIG. 8. The reversible electronic conductivity of a specimen of zinc oxide that had previously been brought to equilibrium at about 1000°C. The measurements were made so rapidly that oxygen did not diffuse in or out of the specimen during the time of measurement. (After P. H. Miller.)

the salt is heated to the vicinity of 1000°C and is then cooled again, two striking changes are found to have occurred, in addition to crystal growth. In the first place, the white material obtains a reddish tint and in the second, it becomes an electronic conductor.

The dependence of electronic conductivity upon temperature for a typical specimen is shown in Fig. 8. The measurements were made so rapidly that the conductivity was reversible. It may be seen that the conductivity obeys a relation of the form given in Eq. (3) for the range of temperature extending from 400°C to 1000°C. If the temperature is maintained at a fixed value, the conductivity varies until it reaches a limiting, or saturation, value that depends upon the pressure of oxygen surrounding the specimen. Figure 9 shows[5] the manner in which this limiting conductivity depends upon the temperature and upon the oxygen pressure. It may be observed that the electronic conductivity decreases with increasing pressure of oxygen.

[5] P. H. Miller, Phys. Rev. 60, 890 (1941).

Careful analysis shows that this behavior is consistent with the following sequence of reactions:

$$2Zn^{++}+2O^{--}\rightleftarrows 2Zn \text{ (in lattice)} + O_2 \text{ (in vapor)}, \quad (5)$$

$$2Zn \text{ (in lattice)} \rightleftarrows 2Zn^{+} \text{ (in lattice)} + 2 \text{ electrons}.$$

In other words, when the material is heated, a partial decomposition occurs and oxygen is liberated. The stoichiometric excess of zinc is first formed at the surface but then diffuses into the interstices of the zinc-oxide lattice, which is very porous on an atomic scale (Fig. 3c). Some of the interstitial zinc atoms become ionized as a result of thermal excitation, and the free electrons are responsible for the electronic conductivity. The reaction is reversible in the sense that the density of interstitial zinc atoms and the electronic conductivity may be reduced by raising the pressure of oxygen.

Semi-conductors of this type, in which the conductivity is due to free electrons and is enhanced by lowering the pressure of the electronegative constituent, are called *electron-excess semi-conductors*.

It should be added that zinc sulfide resembles zinc oxide very closely in that it has a very similar open lattice structure and may be decomposed with the evolution of sulfur by being heated. It turns out that in this case the interstitial zinc atoms that are produced by the heating process are not as readily ionized, so that practically no electrons are free at room temperature. They may be freed, however, by irradiating the crystal with suitable wavelengths of light. As a result, the crystals of zinc sulfide are *photo-conducting*. The interstitial atoms also play an important role in the process

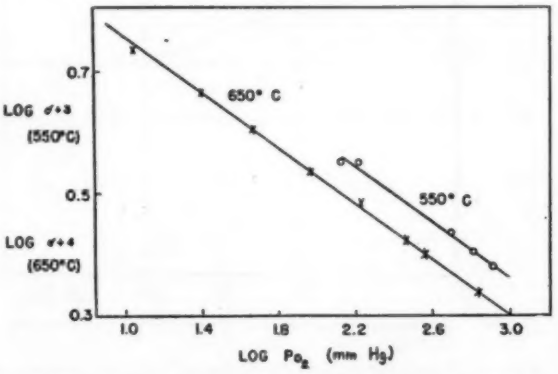


FIG. 9. The dependence of the equilibrium conductivity of zinc oxide upon the pressure of oxygen. (After Wagner.)

by which these crystals are made luminescent when exposed to ultraviolet light.

We note from Fig. 8 that the electronic conductivity of zinc oxide deviates from the relationship given by Eq. (3) when the conductivity it lowered below about 400°C. There is good reason for believing that above this temperature the free electrons are derived principally from isolated monatomic zinc atoms that are bound in the interstices of the lattice. These electrons may then be regarded as constituting a gas or a vapor derived from the neutral zinc atoms. The vapor pressure of this gas falls with decreasing temperature. Now, in addition to the monatomic zinc atoms, there should be present in the lattice neutral, diatomic zinc molecules. Only a small fraction of the stoichiometric excess zinc atoms should actually be in the form of diatomic molecules at the concentrations ordinarily achieved; namely, about 10^{18} atoms per cc. However, it may be shown that the energy required to remove an electron from a diatomic molecule of zinc should be considerably smaller than that required to ionize an isolated atom. As a result, all of the diatomic molecules are ionized at 400°C. For this reason the vapor pressure of the electron gas arising from the diatomic molecules is comparable with that of the gas arising from the monatomic single atoms at that temperature. The molecules remain ionized until room temperature is reached, whereupon, the electronic conductivity again decreases.

This example serves to show that the electronic conductivity derived from a stoichiometric excess of one of the constituents of a compound is greatly affected by the state of aggregation of the atoms of the excess constituent. If all of the excess atoms are present in an atomically dispersed form, the relationship between conductivity and temperature may be relatively simple. If, however, extensive clustering occurs, it may be much more complicated. For this reason, the "aging" processes to which commercial semi-conductors are commonly placed may usually serve to complicate their behavior.

6. COPPER OXIDE[6]

Cuprous oxide is widely used as a commercial semi-conductor and possesses properties that stand in distinct contrast with those of zinc oxide. The reversible conductivity of a typical specimen of this material is shown in Fig. 10. It may be seen that as in the case of zinc oxide, the electronic conductivity varies very rapidly with temperature above a fairly well defined temperature (400°C in this case), and is nearly constant just above room temperature. As in the case of zinc oxide, we may conclude that the atomic centers which are responsible for the free electrons are present principally in monatomic form. However, a relatively small fraction are present in clusters. The monatomic units have a higher heat of ionization than the clustered units, so that the vapor pressure of the conducting gas derived from the latter becomes comparable with that of the former at a temperature near 400°C.

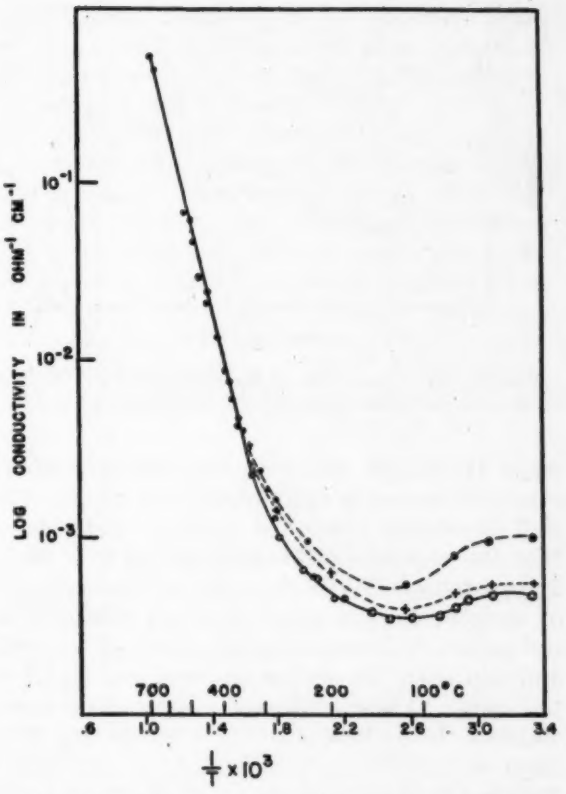


FIG. 10. The conductivity of specimens of cuprous oxide as a function of temperature. (After as yet unpublished work of W. Feldman.)

The differences between the two oxides shows up when we consider the way in which the deviation from stoichiometric proportions occurs. Whereas, the electronic conductivity of zinc oxide falls with increasing oxygen pressure, that of cuprous oxide increases (Fig. 11). This suggests that the electronic conductivity of cuprous

[6] The experiments quoted on cuprous oxide are described more fully in the papers by S. Angello, Phys. Rev. 62, 371 (1942), and W. Feldman, Phys. Rev. 64, 113 (1943).

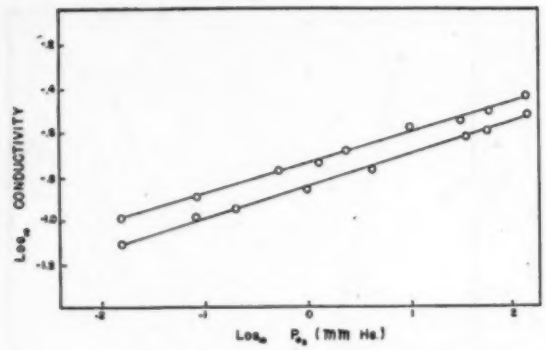


FIG. 11. The dependence of the conductivity of cuprous oxide upon pressure of oxygen. (After Feldman.)

oxide is related either to the presence of an excess of oxygen or to a *deficiency* of copper. The well developed theory of cuprous oxide shows that the sequence of events that occur is probably as follows: When the solid is in the presence of oxygen, oxygen atoms become adsorbed on the surface. Cuprous ions diffuse out of the solid and join with the oxygen to form new layers of the oxide. These diffusing copper ions leave vacancies behind them and take an electron with them, so that there are regions in the solid which possess a deficiency of one atom of copper and a deficiency of one electron. An atom or ion possessing a deficiency of an electron is commonly said to possess a *hole*. It can be shown that these holes may migrate about the lattice and impart an electronic conductivity to it. Thus the electronic conductivity of cuprous oxide arises from the presence of the holes left by the copper atoms which have diffused to the surface. The holes are normally bound to the copper ion-vacancies; however, they may be freed thermally and then behave like free electrons. In this sense the vacancies to which holes are bound are the centers which play a role analogous to the interstitial zinc atoms in zinc oxide.

It is clear from this description that the electronic conductivity of cuprous oxide should increase with increasing oxygen pressure since such an increase in pressure fosters the growth of surface layers and thereby engenders vacancies and holes. The chemical reaction is as follows:

$$4Cu^+ + 2O^{--} + O_2(\text{gas}) \rightleftarrows 8Cu^+ + 4O^{--} + 4\Delta Cu^+ (\text{crystal}) + 4 \text{ holes}, \quad (6)$$

in which ΔCu^+ represents a copper ion-vacancy. Semi-conductors which conduct because of the presence of holes are called *electron-defect* semi-conductors.

It should be added that a semi-conductor containing free holes instead of free electrons possesses several properties which differentiate it from an electron-excess semi-conductor. Most notably, the algebraic signs of its Hall coefficient and its thermoelectric coefficient are the reverse of that found in electron-excess conductors.

7. COPPER-IODIDE

Cuprous iodide is a semi-conductor which resembles copper oxide in practically all respects. It is of particular interest because it absorbs excess iodine at room temperature, so that the changes in composition and in electronic conductivity may be detected with relative ease by continuous simultaneous measurements of weight and conductivity.

8. LEAD CHROMATE[7]

Monoclinic lead chromate, which is extensively used as a commercial yellow pigment, darkens when heated and at the same time develops an electronic conductivity. Recent experimental work by Lashof has shown that the reaction governing this example is probably as follows:

$$2PbCrO_4 \rightleftarrows Pb(CrO_2)_2 \text{ (in solid)} + 2O_2(\text{gas}) + Pb^\circ \text{ (in solid)} \quad Pb^\circ \rightleftarrows Pb^+ + \text{electron}.$$

In other words, the heating process yields oxygen gas, neutral lead atoms, and transforms chromate ions into chromite ions. Since chromate ion is yellow, whereas chromite is green, the change in valency of the chromium should be associated with a darkening of the yellow pigment, as is observed. The production of free lead atoms should transform the material into an electronic semi-conductor of the electron-excess type, as is also observed.

9. MONATOMIC SEMI-CONDUCTORS

As we remarked in the introduction, there are a number of semi-conductors, such as selenium, which are monatomic and which as a result cannot derive their conductivity from deviations from stoichiometric proportions of the constituents. In such cases the conducting electrons or holes arise either from impurities which play the same role as the lattice defects in the salts, or from the atoms of the base material. In the second case, the atoms of the bulk material hold

[7] The experiments quoted on lead chromate are described more fully in the papers by T. W. Lashof, J. Chem. Phys. 11, 196 (1943), and J. Goldman and A. W. Lawson, Phys. Rev. 64, 11 (1934).

their valence electrons at low temperatures but become thermally ionized at elevated temperatures. An electronic conductivity of this type is called *intrinsic*.

SUMMARY

Typical metal alloys exhibit striking deviations from the conventional laws of combining proportions. An investigation of the origin of these deviations leads us to expect that similar deviations will occur even in the most ideal ionic and homopolar compounds at elevated temperatures. Experiments show that such deviations actually do occur, and are usually accompanied by the development of electronic conductivity, which may be used as a means of ascertaining the nature of the chemical changes. The experimental evidence dealing with pure monatomic substances shows that an electronic conductivity may occur even in cases in which deviations from ideal stoichiometric ratios are not present. Thus, the underlying principles which govern the properties of semi-conductors are somewhat broader than those of interest for stoichiometry alone. Nevertheless, the investigation of semi-conductors does provide us with a beautiful confirmation of an extension of the ordinary rules of combining proportions.

The Electrical Properties of Semi-Conductors*

BY ROBERT J. MAURER

Carnegie Institute of Technology, Pittsburgh, Pennsylvania

The electrical properties of semi-conductors are briefly reviewed. The information concerning the structure of these materials which is obtained by electrical measurements is discussed. The most important of the physical-chemical methods available for the investigation of deviations from stoichiometric proportions in semi-conducting compounds are illustrated by a discussion of several typical examples.

INTRODUCTION

SEMI-CONDUCTORS are a class of solids whose electronic conductivity falls in the range 10^{-5} to $10^{3}\Omega^{-1}-cm^{-1}$. This is to be compared with the typical conductivities of metals which are of the order of $10^{5}\Omega^{-1}-cm^{-1}$. The significant difference between semi-conductors and metals is not, however, the difference in the order of magnitude of their conductivities but the tremendous variation of the conductivity of semi-conductors with the chemical and thermal treatment given them.

Our knowledge of the fundamental electrical properties of semi-conductors is small by comparison with that of metals. The quantum mechanical theory of solids has progressed to a point where it is possible to make a quantitative comparison of the calculated and measured conductivities of a specific metal. In discussing semi-conductors we must still be content with general models. It is the purpose of this paper to review some of the important experiments which have yielded information regarding the behavior of electrons in semi-conductors. Because of the peculiar nature of the origin of the conductivity of these materials, it is impossible to do this and neglect their chemistry.

ELECTRICAL CONDUCTIVITY

Our introductory discussion has tacitly assumed that the conductivity of a semi-conductor is its most important electrical property. Figure 1 shows a typical experimental arrangement for the measurement of conductivity and its temperature variation. The potential drop across a portion of the specimen is measured with a known current passing through it. The purpose of this method is to avoid errors due to a large contact resistance which is frequently present at the junction of a metal and a semi-conductor. It must be recognized that the conductivity may not be due entirely to electrons. The ionic conductivity may be very small, as in cuprous oxide at 1000°C, where it amounts to only 4×10^{-4} of the total conductivity.[1] On the other

* Based on a paper presented at a symposium on semi-conductors at the Detroit meeting of the American Chemical Society, April, 1943.

[1] J. Gunderman and C. Wagner, Zeits. f. physik. Chemie B37, 155 (1937).

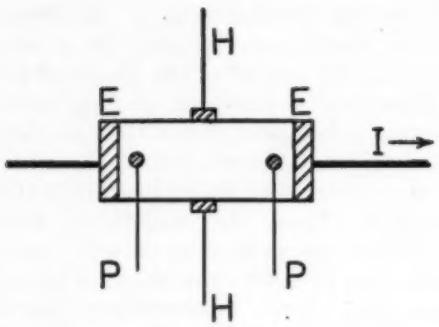


FIG. 1. Measurement of conductivity and Hall effect; E,E are electrodes for passing the constant current, I, through sample, P,P and H,H are electrodes for potentiometric measurement of voltage drop and Hall e.m.f., respectively. Magnetic field is assumed perpendicular to plane of paper.

hand, the electrolytic conductivity of β-Ag_2S at 170°C is 81 percent of the total conductivity.[2]

The electronic conductivity of a semi-conductor can be expressed in terms of an equation which was originally proposed by Lorentz[3] to describe the conductivity of metals.

$$\sigma = \frac{4}{3}\frac{e^2 n l}{(2\pi m k T)^{\frac{1}{2}}}, \tag{1}$$

where e is the electronic charge, k is Boltzmann's constant, T is the absolute temperature, n is the density of conducting electrons, and l is their mean free path. The density of the free electron gas in a metal is too large to permit the use of the Maxwell-Boltzmann statistics as was done by Lorentz in the derivation of this equation. In semi-conductors, the density of conducting electrons is seldom greater than $10^{18}/cm^3$ and Lorentz' equation is valid. The mass, m, which appears in the equation is an effective mass which may be greater or less than the electronic mass. It will be noted that the product of n and l appears in (1). The determination of the conductivity fixes neither of these important quantities uniquely.

Over rather wide ranges of temperature the conductivity of a semi-conductor can usually be expressed by an empirical equation of the form

$$\sigma = Ae^{-E/kT}, \tag{2}$$

where A and E are constants. This means that either the density of free electrons or the mean free path, or both, are strongly temperature dependent. It has been frequently assumed in the past that the mean free path of an electron in a semi-conductor is determined by the same mechanisms as in a metal. In a metal the most important factor is the interaction of the electron with the ions of the lattice as a result of the thermal motion of the latter. This results in a transfer of the energy, which the electron has gained from the applied electric field, to the lattice with the stimulation of lattice vibrations. At ordinary and high temperatures it is well known that this process leads to the approximately linear increase of resistance of metals with temperature. The exponential increase of the conductivity of semi-conductors with temperature has been ascribed to an exponential increase in the number of free electrons. Furthermore, since the presence of traces of impurities in metals leads to an increase in resistance which is small and independent of the temperature, the great dependence of the resistance of semi-conductors upon their impurity content has been ascribed to the effect of these impurities upon the number of free electrons. In fact, it is believed that in the pure state at ordinary temperatures practically all semi-conductors are good insulators. The impurities are a source of potentially free electrons. To free an electron from an impurity center requires an activation energy, E, and the density of free electrons follows from Boltzmann's principle

$$n = n_0 e^{-E/kT}. \tag{3}$$

Wilson[4] has been able to fit these ideas into the quantum-mechanical theory of solids. We shall not pursue this phase of the topic but will point out one important result. An impurity center may possess no bound electrons which can be thermally freed but may be able to trap one of the "normal" electrons of the semi-conductor. This leaves a "hole" in the electron distribution of the solid. Such "holes" will give rise to a conductivity described by Eqs. (1)–(3) but the sign of the charge carriers will appear to be positive. This peculiar "hole" conduction, which occurs in many metals and is predicted by the theoretical treatment of the motion of an electron in the periodic potential field of a lattice, can be detected by measurement of the Hall effect.

[2] C. Tubandt and H. Reinhold, Zeits. f. anorg. allgem. Chemie **160**, 222 (1927).

[3] F. Seitz, *Modern Theory of Solids* (McGraw-Hill Book Company, Inc., New York, 1940), p. 190.

[4] A. H. Wilson, *Semi-conductors and Metals* (Cambridge University Press, New York, 1939).

THE HALL EFFECT

When an electronic current is present in a material and there is applied a uniform magnetic field in a direction perpendicular to that of the current, an e.m.f. appears across the sample in a direction perpendicular to the plane defined by the current and magnetic field vectors. This e.m.f. is proportional to the product of the current and the magnetic field strength and inversely proportional to the thickness t of the specimen.

$$\epsilon = R(Hi/t). \tag{4}$$

An extension[5] of Lorentz' treatment of an electron gas to the case where electric and magnetic fields are present gives the following equation for the constant R:

$$R = (-3\pi/8)(1/nec). \tag{5}$$

The measurement of the Hall constant, R, yields information of two kinds. The sign of R tells us whether the conduction is due to electrons or to "holes." The numerical magnitude of R allows the calculation of the density of conducting electrons or "holes." If the Hall constant and the conductivity are measured the mean free path can be obtained from Eqs. (1) and (5)

$$l = \left(\frac{8c^2mkT}{\pi e^2}\right)^{\frac{1}{2}} \sigma R. \tag{6}$$

The experimental arrangement for the measurement of the Hall e.m.f. is shown in Fig. 1. The Hall e.m.f. is not easy to determine because it is usually of the order of a fraction of a millivolt and at high temperatures is masked by thermoelectric e.m.f.'s which occur in the circuit as a result of unavoidable thermal gradients across the sample.

Figure 2 shows the mean free path in cuprous oxide at low temperatures. These results are calculated by Engelhard[6] from Hall effect and conductivity data taken by him. The experimental temperature dependence agrees well with the theory of Fröhlich and Mott[7] which predicts the rapid increase of the mean free path at low temperatures.

Miller's data upon the conductivity of ZnO have been described by Seitz[8] in the preceding paper. The conduction electrons result from the thermal excitation of electrons from interstitial zinc atoms. In Miller's work the experiments were conducted in a time short compared with that necessary for equilibrium to be established so that the number of interstitial zinc atoms was kept constant. Miller used the slope of the high temperature portion of the curve to determine the activation energy for the release of an electron from an interstitial zinc atom with the assumption of a temperature independent mean free path. The change in slope of the low temperature part of the curve he ascribed to the pairing of zinc atoms with the production of interstitial zinc molecules. On purely statistical grounds there will be a small fraction of these pairs present. Since the activation energy for the pairs will be lower than that of the atoms, the behavior of the pairs will dominate the low temperature part of the curve where the number of electrons released from atoms will be small as compared with the number released from the pairs. Figure 3 gives the data of Miller[9] on the

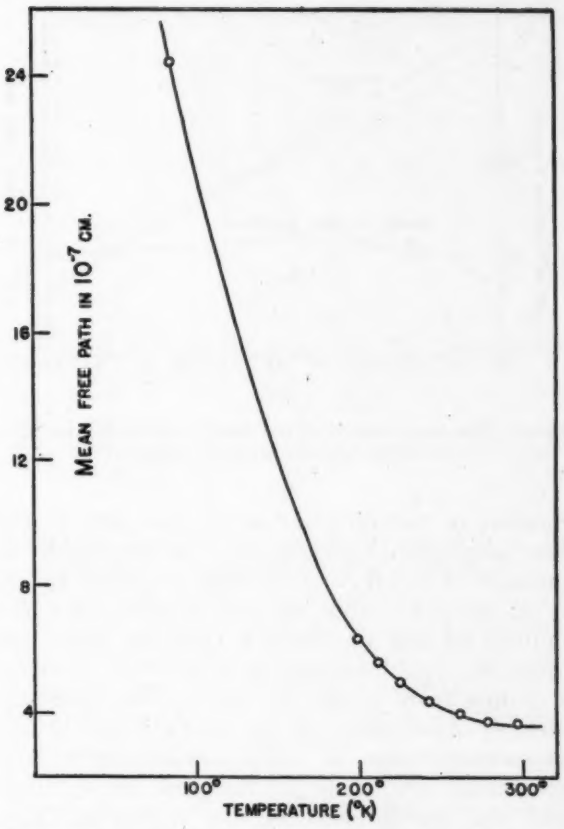


FIG. 2. The mean free path in cuprous oxide as a function of temperature (after Engelhard).

[5] F. Seitz, see reference 3.
[6] E. Engelhard, Ann. d. Physik **17**, 501 (1933).
[7] H. Fröhlich and N. Mott, Proc. Roy. Soc. **A171**, 496 (1939).
[8] F. Seitz, J. App. Phys. **16**, 553 (1945).
[9] P. Miller, Jr., Phys. Rev. **60**, 890 (1941).

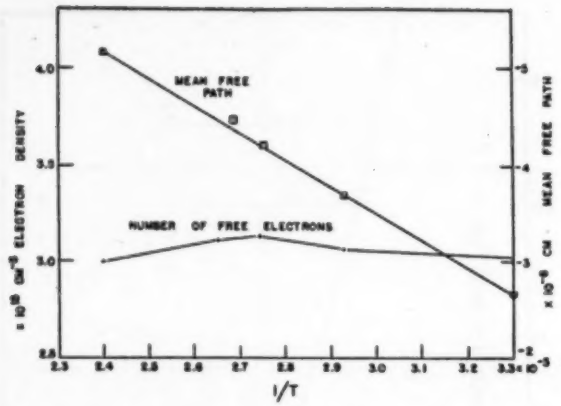


FIG. 3. The mean free path and density of conduction electrons in zinc oxide (after Miller).

number of electrons and mean free path in the low temperature region. It was impossible to measure the Hall e.m.f. at high temperatures. It is to be noted that at low temperatures the number of free electrons is constant while the mean free path increases by a factor of two over a temperature range of 100°C. The constant number of free electrons can be explained by the assumption that at these temperatures each interstitial zinc molecule has lost an electron, and the contribution of the interstitial zinc atoms is still negligible. The number of molecules, determined from the electrical measurements in this manner, is what one would expect from the density of interstitial zinc atoms and the statistical probability of a certain fraction being nearest neighbors in the lattice. The rapid change in mean free path is anomalous. This behavior is not unique, however. Feldman[10] found that the mean free path in cuprous oxide was a strong function of the temperature in the neighborhood of 300°C. It must be realized that in the interpretation of conductivity data the assumption of a mean free path which is independent of temperature and concentration of impurity is extremely dangerous.

THE THERMOELECTRIC EFFECTS

The thermoelectric effects in solids are rather complicated second-order effects so that little precise information can be obtained from them. The sign of the Seebach e.m.f. gives, however, the sign of the charges whose motion constitutes the current and its determination is a relatively simple matter. It is known that the reason for the fact that the Seebach e.m.f. at a semi-conductor-metal junction is a thousand times greater than at metal-metal junctions is the result of the great difference in the density of conduction electrons in semi-conductor and metal but differences cannot be determined precisely in this manner.

DEVIATIONS FROM STOICHIOMETRIC PROPORTIONS

The experiments which have been described have all been performed upon semi-conductors which possessed an unknown constant deviation from stoichiometric proportions. Obviously, one of the most important problems is the correlation of the purely electrical properties with the excess or deficiency of a given component in the compound. In addition there is the purely chemical problem of determining how the extent of the deviation from stoichiometric proportions depends upon the equilibrium conditions.

The types of deviations from stoichiometric proportions which are to be expected in semi-conducting compounds have been reviewed by Seitz.[8] First, there is the wandering of one type of ion into an interstitial position while the other type of ion leaves the crystal for the gas phase. This is usually referred to as Frenkel disorder. Another type of disorder is the Schottky type where the ions wander to the surface and build an extension of the lattice with the simultaneous formation of vacancies. In principle, the type of disorder can be established by purely experimental methods. The combination of density measurement and x-ray determination of the lattice constant gives the number of molecules per unit cell,

$$\nu = Na^3d/M,$$

where ν is the number of molecules per unit cell, N is Avogadro's number, a is the lattice constant, d is the density of the sample, and M is the molecular weight. In the case of Frenkel disorder, the average number of molecules per unit cell becomes larger as the degree of disorder increases. With Schottky disorder the density of molecules decreases as the degree of disorder increases. The number of vacancies per unit volume is usually in the range from 10^{-7} to 10^{-4} so that the precision of the measurements is ordinarily insufficient to make the method of practical value. The Wüstite phase of FeO is an exception. Wüstite has the sodium chloride type of lattice, and the iron content can be

[10] W. Feldman, Thesis, University of Pennsylvania (1942).

varied from 76.08 to 76.72 wt. percent of iron. Jette and Foote[11] prepared a large number of samples with different iron contents and determined the iron content by chemical analysis. Density and x-ray measurements then enabled them to show that the disorder was of the Schottky type. One iron ion diffuses to the surface for each excess oxygen atom present in the solid phase. The oxygen atom becomes an ion by obtaining two electrons from the solid. The result is the creation of iron ion vacancies and electron "holes." In this case the degree of disorder is so great that it is measurable by chemical analysis.

Juse and Kurtschatow[12] were able to demonstrate the presence of excess oxygen in cuprous oxide by chemical analysis. The cuprous oxide was dissolved in acid potassium iodide solution and the oxidized iodine titrated with hyposulfite solution. Samples of cuprous oxide which had been heated in a vacuum showed no detectable excess oxygen. Samples heated in oxygen possessed excess oxygen which increased with the time of heating and reduction of the annealing temperature. Their results are tabulated in Table I. These authors determined also the conductivity *versus* temperature curves for their samples. The activation energies were the same for all of the samples possessing excess oxygen, but the magnitude of the constant A in Eq. (2) increased with the amount of excess oxygen. Dünwald and Wagner[13] also attempted to determine the amount of excess oxygen in cuprous oxide by determining the change in oxygen pressure in a sealed flask when the oxygen was exposed to the solid. They established that the amount of oxygen absorbed varied approximately as the seventh root of the oxygen pressure,

TABLE I. Excess oxygen in cuprous oxide as determined by chemical analysis. (Juse and Kurtschatow.)

Remarks	Excess oxygen in wt. percent
Original material	0.053
Vacuum heated	not detectable
1 hour in air at 1000°C	0.060
6 hours in air at 700°C	0.084
6 hours in air at 600°C	0.090
6 hours in air at 500°C	0.101

[11] E. Jette and F. Foote, J. Chem. Phys. 1, 29 (1933).
[12] W. Juse and B. Kurtschatow, Physik. Zeits. Sowjetunion 2, 453 (1932).
[13] H. Dünwald and C. Wagner, Zeits. f. physik. Chemie B22, 212 (1933).

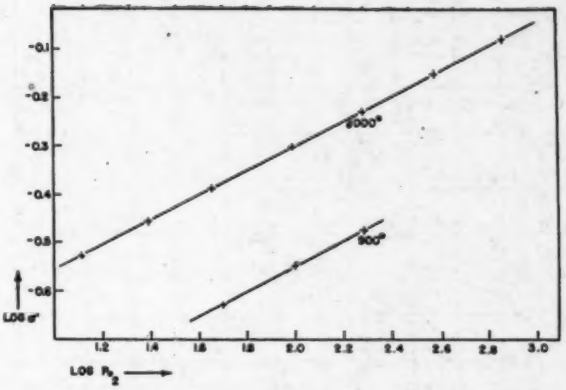


FIG. 4. The logarithm of the conductivity of nickelous oxide *versus* the logarithm of the oxygen pressure (after Baumbach and Wagner).

which is the same pressure dependence as the conductivity. Baumbach and Wagner[14] attempted to detect a change in weight of nickelous oxide with change of oxygen pressure at 1000°C. Any change which occurred was less than that corresponding to an absorption of 2×10^{-3} g-atom of O per mol of NiO, which was the limit of sensitivity of their apparatus.

The dependence of the conductivity upon the vapor pressure of the gas phase of one component of a semi-conducting compound has frequently been used to give information regarding the course of the reaction which results in the production of disorder. Several examples have been given by Seitz in the preceding paper.

Another example is given in Fig. 4 which shows the dependence of the conductivity of nickelous oxide upon the pressure of oxygen in equilibrium with the solid.[14] The conductivity varies as the fourth root of the oxygen pressure. A mechanism which will yield this behavior is expressed in the following equation:

$$Ni^{++}+O^{=}+\tfrac{1}{2}O_2(g)\rightleftharpoons Ni^{++}+2O^{=}+[Ni^{++}]^{+}+h.$$

The reaction can be described as the diffusion of a nickelous ion to the surface where it builds an extension of the crystal lattice with the atom of oxygen from the gas phase. Two electrons diffuse to the surface to convert the oxygen atom into a double charged ion. One of the two "holes" which result is trapped at the vacancy left by the nickelous ion in the lattice and the symbol $[Ni^{++}]^{+}$ indicates this vacancy plus trapped "hole." The other "hole" is free and is

[14] H. Baumbach and C. Wagner, Zeits. f. physik. Chemie B22, 59 (1934).

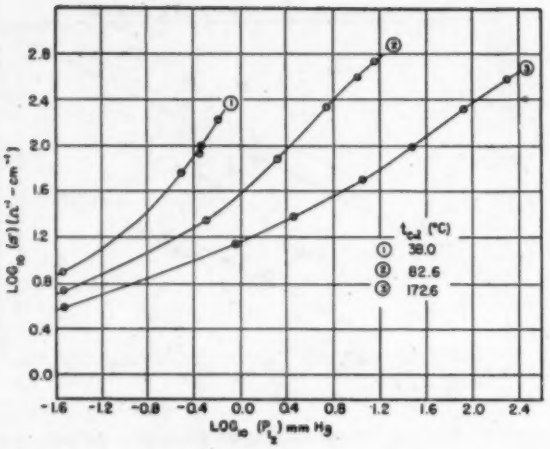


FIG. 5. The logarithm of the conductivity of cuprous iodide *versus* the logarithm of the iodine pressure.

the source of the electrical conductivity. The application of the mass action law to the above equation and the assumption that the conductivity is proportional to the concentration of holes leads to the observed dependence of the conductivity on the pressure.

When the conductivity does not vary as some simple power of the pressure of the gas in equilibrium with the solid, it is usually impossible to deduce the mechanism of the reaction from the pressure-conductivity curves. As an example of this type of behavior, we may refer to cuprous oxide below 800°C; and to cuprous iodide.

The conductivity of cuprous oxide varies according to the following equation:

$$\sigma = \sigma_0 p_{O_2}^{1/7}$$

for temperatures between 1000°C and 800°C, while the theory predicts an eighth root dependence upon the pressure of oxygen. Below 800°C there is no simple relationship between pressure and conductivity.

Cuprous iodide is a semi-conductor whose electrical properties depend upon the amount of excess iodine in a very complicated fashion. Nagel and Wagner[15] have proposed the following reaction:

$$Cu^+ + I^- + \tfrac{1}{2}I_2(g) \rightleftharpoons Cu^+ + 2I^- + [Cu^+] + h.$$

$[Cu^+]$ denotes a cuprous ion vacancy and h an electron hole. A cuprous ion and an electron are supposed to wander to the surface and combine with an iodine atom from the gas phase. The result is the formation of cuprous ion vacancies and holes. At a fixed temperature, the fraction of the holes which are free is assumed constant. An application of the mass action law to the above reaction, in which the vacancies and holes are treated as chemical entities, then leads to a fourth root dependence of the conductivity upon the pressure of the iodine vapor in equilibrium with cuprous iodide.

$$\sigma = \sigma_0 p_{I_2}^{1/4}.$$

It is assumed that the mean free path is independent of and the conductivity proportional to the density of holes. Figure 5 shows the experimental dependence of the conductivity upon the pressure of iodine vapor.[16] There is no simple power law relation between conductivity and pressure.

The weight of excess iodine absorbed by cuprous iodide varies, however, as the square root of the iodine pressure as shown in Fig. 6. This data was obtained with the quartz microbalance illustrated in Fig. 7. Magnetic control of the balance in the manner described by Blewett[17] allowed the balance to be operated in an atmosphere of iodine. The increase in weight of the sample due to absorption of iodine was balanced by the torque upon the iron cylinder caused by the magnetic field of a pair of Helmholtz coils. The iron cylinder and platinum counterweight

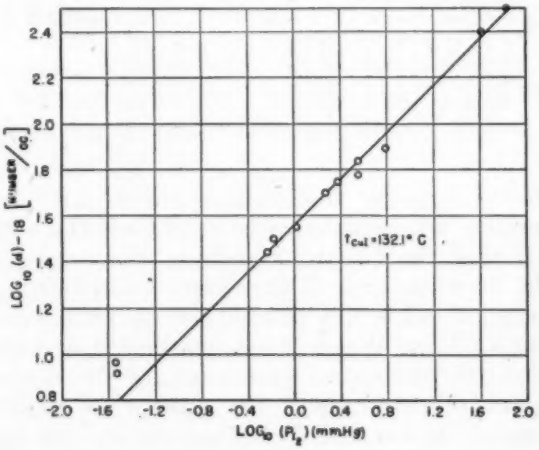


FIG. 6. The logarithm of the density of absorbed iodine atoms *versus* the logarithm of the iodine pressure.

[15] K. Nagel and C. Wagner, Zeits. f. physik. Chemie **B25**, 71 (1934).

[16] This work was supported by the Solar Energy Research Fund of Massachusetts Institute of Technology and carried out in the laboratory of Professor A. von Hippel. A detailed account will be published soon.

[17] J. P. Blewett, Rev. Sci. Inst. 10, 231 (1939).

were sealed in quartz to protect them from the iodine vapor.

That the dependence of the conductivity of cuprous iodide upon the amount of excess iodine is very complicated is further illustrated by Fig. 8. Here we have the mean free path as a function of the density of holes as calculated from conductivity and Hall effect data of Steinberg.[18] The mean free path drops rapidly as the density of holes becomes greater than 10^{18} cc^{-1}.

The alkali halide semi-conductors have been extensively investigated by Pohl and his coworkers who have developed a very interesting optical method for the determination of the amount of excess alkali metal present. The reaction proposed for the absorption of excess potassium in potassium chloride is

$$K^+ + Cl^- + K(g) \rightleftharpoons 2K^+ + Cl^- + [Cl^-]^-.$$

When the crystal is heated at a high temperature in the presence of potassium vapor chlorine ions wander to the surface of the crystal and build on to the lattice with potassium atoms from the vapor. The potassium atoms lose electrons which diffuse into the crystal and are trapped at the vacant lattice sites where chlorine ions are missing. These sites will be the most stable positions in the crystal for the electrons. An electron trapped at a chlorine vacancy can absorb light and be excited to a higher energy level. The absorption band corresponding to this process is shown in Fig. 9.

The reason for the presence of an absorption band rather than an absorption line is the interaction of the electron with the lattice. If the half-width and height of the absorption band are measured, the number of absorbing centers

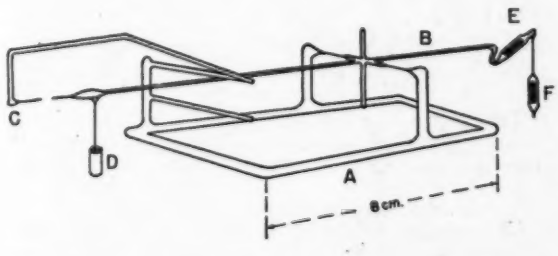


FIG. 7. Quartz microbalance for operation in an atmosphere of iodine. *A*, frame; *B*, beam; *C*, fiducial pointer; *D*, sample; *E*, iron cylinder, quartz enclosed; *F*, quartz enclosed platinum counterweight.

[18] K. Steinberg, Ann. d. Physik **35**, 1009 (1911).

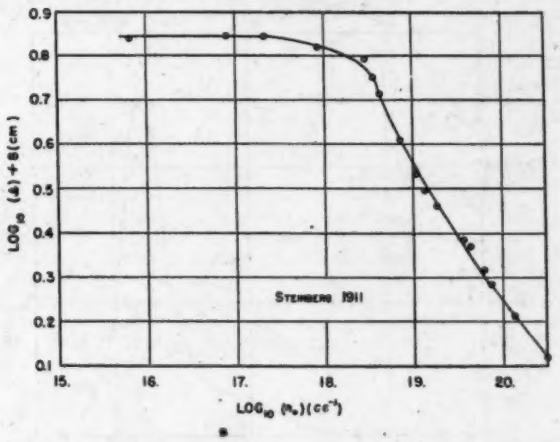


FIG. 8. The logarithm of the mean free path in cuprous iodide *versus* the logarithm of the density of conduction "holes" (after Steinberg).

can be calculated from dispersion theory.[19] There is one unknown factor, the oscillator strength which is proportional to the probability that a trapped electron will absorb a photon, but there are very good theoretical grounds for believing that this strength is close to unity. Kleinschrod[20] compared the amount of excess potassium in potassium chloride crystals as determined by the optical method and chemical analysis. He dissolved his crystals in water and determined the resultant *p*H of the solution using methyl red as an indicator. The absorption curves of the methyl red were determined before and after solution of the potassium chloride and the excess potassium obtained from the change in *p*H. For concentrations of excess potassium atoms between $1.0\times10^{+17}$ and $2.0\times10^{+18}$ cc^{-1}, the chemical procedure gave an increase of 19 percent over the concentrations determined optically. The interpretation is that the oscillator strength of an electron trapped at a halogen ion vacancy is not 1.0 but 0.8. The usefulness of the optical determination depends upon the occurrence of the absorption band due to the vacancies in an accessible portion of the spectrum where it does not overlap the characteristic absorption of the pure crystal. In addition, it is necessary that the absorption band have the simple bell shape predicted by optical dispersion theory.

CONCLUSION

Deviations from stoichiometric proportions are the most important type of "impurity" which

[19] A. Smakula, Zeits. f. Physik **59**, 603 (1930).
[20] F. Kleinschrod, Ann. d. Physik **27**, 97 (1936).

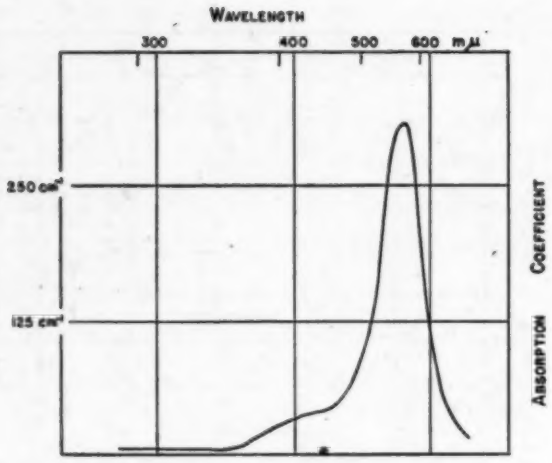


FIG. 9. Absorption coefficient *versus* wave-length for trapped electrons in KCl (after Pohl).

one has to consider when dealing with the electrical properties of semi-conducting compounds There are available physical-chemical methods by which, in favorable cases, the degree of disorder may be determined as a function of the equilibrium conditions. The combination of these measurements with the determination of the electrical conductivity, σ, and the Hall constant, R, enables one to deduce the mean free path and the density of conducting particles as a function of the degree of disorder and of the temperature. The absolute value of the mean free path is not uniquely fixed by these measurements because there is no reliable method of determining the effective mass of the particles, which occurs in Eq. 4 for the mean free path. It is assumed, ordinarily, that the effective mass is the electronic mass.

On the Representation of Rigid Rotations

BY H. SCHWERDTFEGER

Department of Mathematics, The University of Adelaide, Adelaide, Australia

This expository article gives a straight-forward derivation of the classical representation formulas of the three-dimensional rotation group in explicit (non-symbolical) form starting with the most elementary representation of a rotation by a linear homogeneous transformation with a proper orthogonal matrix. Apart from Euler's formula (3.4) (which is taken as starting point in Benedikt's previous paper on the same subject) other representations often used in mechanics and physics are derived and the various connections existing between these formulas are established hereby going beyond Benedikt's paper by taking into account the simplest spin representations of the rotation group. The spin representation by the special unitary group is derived in two ways; the second method (Section 6) is based on a general fact (in italics) which seems not to be generally known. In the footnotes some references to the literature are given.

IN a recent note E. T. Benedikt[1] discusses the classical parametrization of the rotation group in three dimensions. Taking up this subject again, it is the purpose of the present note to derive the fundamental formula by systematic application of matrix calculus[2] and to show how from this point of view its connection with other forms of the representation of the rotation group may easily be established.

1. NOTATIONS

A vector in the three-dimensional space will be represented by a column matrix

$$x=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix},\quad p=\begin{pmatrix}p_1\\p_2\\p_3\end{pmatrix},\quad \cdots.$$

We associate with the column x

(i) the row matrix $x'=(x_1\,x_2\,x_3)$, (transposition),
(ii) the skew-symmetric matrix

$$[x]=X=\begin{pmatrix}0&-x_3&x_2\\x_3&0&-x_1\\-x_2&x_1&0\end{pmatrix},\quad X'=-X.$$

By means of these "elements" and using the fundamental rules of matrix algebra, one obtains all formulae of the orthogonal vector algebra (between these several formulae not appearing in the usual treatises) as simple matrix identities.[3] In this way the column Xy represents the vector product of the two vectors x, y; the associated skew-symmetric matrix is

$$[Xy]=XY-YX=yx'-xy', \qquad (1.1)$$

where xy' is the column-row product (representing the often so-called "dyadic product" of the two vectors x, y), i.e., the three-rowed square matrix with the elements $x_\mu y_\nu$. The inner product of x and y is represented by $x'y$. Further, one has the double vector product

$$XYz=x'z\cdot y-x'y\cdot z, \qquad (1.2)$$

whence it follows the fundamental identity between any two skew-symmetric matrices

$$YXY=-x'y\cdot Y, \qquad (1.3)$$

which for $y=x$ yields Cayley's identity of X, i.e.,

$$X^3=-x'x\cdot X. \qquad (1.4)$$

By using in (1.2) the three unit columns successsively instead of z, one obtains the following relation

$$XY=yx'-x'y\cdot E, \qquad (1.5)$$

where E denotes the unit matrix. Finally we notice the determinant formula

$$|x\,y\,z|=x'Yz. \qquad (1.6)$$

2. THE ROTATION GROUP

A rotation in space about the fixed origin 0 of the vector x is given analytically by a linear

[1] E. T. Benedikt, J. App. Phys. **15**, 613–615 (1944). A recent book covering our present subject in greater generality is F. D. Murnaghan, *The Theory of Group Representations* (Baltimore, 1938).

[2] In this form the matter in question has been dealt with by the author as an example in his lectures on Linear Algebra and Matrix Calculus at the University of Adelaide during recent years. Therefore the reader of this note is supposed to be conversant with the more elementary parts of this topic as it is found in various modern algebraic texts or in the exposition by L. A. Pipes, "Matrices in Engineering," Elec. Eng. **56**, 1177–1190 (1937); cf. also H. Schwerdtfeger, *Les Fonctions de Matrices I, Actualités scientifiques* (Paris, 1938), p. 649.

[3] For further details on this subect see the author's booklet quoted in reference 2, Chap. I, Section 8; cf. also H. Schwerdtfeger, "Skew-symmetric matrices and projective geometry," Am. Math. Mon. **51**, 137–148 (1944).

transformation $\bar{x}=Ax$ where A is a proper orthogonal matrix, i.e., a real matrix for which

$$A'A=E \quad \text{and} \quad |A|=1. \tag{2.1}$$

The set of all rotations forms a group which is isomorphic with the group **G** of all orthogonal matrices with positive determinant when the group composition is the matrix multiplication.

Now let $P=[p]$ be a real skew-symmetric matrix. Then the matrix $A=e^P$ (the matrix e^P being defined by the power series of the exponential function) satisfies the conditions (2.1). In fact

$$A'A=\exp(P')e^P=\exp(P'+P)=e^0=E.^4$$

Further let

$$p'p=\omega^2;$$

it follows from (1.4) that 0, $i\omega$, $-i\omega$ are the three latent roots (also called the characteristic values or "eigenvalues") of P; therefore

$$|e^P|=e^0\cdot e^{i\omega}\cdot e^{-i\omega}=1.$$

If t is a real parameter, usefully interpreted as time scale of the motion, evidently the matrices

$$B(t)=e^{Pt}$$

form a one-parameter subgroup $\mathbf{G}_1$ of **G**, *viz.*,

$$B(t_1)\cdot B(t_2)=B(t_1+t_2),$$

which contains the above matrix $e^P=B(1)$.

Conversely let A be a fixed proper orthogonal matrix. We shall show that A then is an element of a certain one-parameter subgroup $\mathbf{G}_1$ of **G**. Consider a vector x of length one and the unit circle about 0 through x and the point $\bar{x}=Ax$. All rotations B such that Bx is a point of this circle form a group $\mathbf{G}_1$ containing A. To prove this we may choose the rectangular coordinate system such that $x=\begin{pmatrix}1\\0\\0\end{pmatrix}$ and the circle is in the x_1x_2-plane: $\bar{x}=\begin{pmatrix}\cos\theta\\ \sin\theta\\ 0\end{pmatrix}$. Then B has necessarily the form $\begin{pmatrix}\cos t & -\sin t & 0\\ \sin t & \cos t & 0\\ 0 & 0 & 1\end{pmatrix}$ and all these matrices are readily seen to form a group $\mathbf{G}_1$ containing $A=B(\theta)$. This group is uniquely defined by A if x and $\bar{x}$ are not diametrally opposite points of the unit sphere.

[4] Notice that $P'P=PP'=-P^2$. Cf. the remark at the end of this section, or R. A. Frazer, W. J. Duncan, and A. R. Collar, *Elementary Matrices* (Cambridge, 1938), p. 42.

By differentiating B with respect to the time parameter t we obtain easily for B the matrix differential equation

$$\dot{B}=QB, \tag{2.2}$$

where Q is the skew-symmetric matrix $\begin{pmatrix}0&-1&0\\1&0&0\\0&0&0\end{pmatrix}$ which, evidently, is associated with the vector of the positive normal of the plane (or circle) defining the subgroup $\mathbf{G}_1$, since with respect to the chosen coordinate system the column $\begin{pmatrix}1\\0\\0\end{pmatrix}$ represents this vector. This geometrical fact remains unaltered if the coordinate system is turned into an arbitrarily fixed position. Hence every one-parameter subgroup $\mathbf{G}_1$ of **G** is defined by a matrix differential Eq. (2.2) where Q is the skew-symmetric matrix associated with the unit-vector of the common axis of all rotations of the subgroup. By integration of (2.2) for the initial condition $B(0)=E$ one obtains

$$B=e^{Q\cdot t}, \tag{2.3}$$

and thus[5a]

$$A=e^{Q\theta}=e^P=E+\sum_{\nu=1}^{\infty}\frac{1}{\nu!}P^\nu, \tag{2.4}$$

where $P=\theta Q$ is a certain skew-symmetric matrix the vector p of which has the direction of the axis of the rotation; its components p_1, p_2, p_3 are the so-called "canonical parameters" of the rotation group. The infinite series is readily shown to be convergent for every matrix A.

In spite of the symbolically simple form of the representation (2.4) it does not easily lend itself, so far, to applications where the analytic expressions of several rotations are to be composed. In fact, a practicable composition rule is given by the functional equation $\exp(P^{(1)})\exp(P^{(2)})=\exp(P^{(1)}+P^{(2)})$ of the exponential function *if, and only if,* $\exp(P^{(1)})$ and $\exp(P^{(2)})$ are elements of the same one-parameter subgroup, i.e., if both rotations have the same axis. In general a simple composition rule in the canonical parameters is not known.[5b]

[5] (a) This is the analog of Benedikt's symbolical formula (10) or (10'). Like our formula (2.4) it is obvious from the point of view of Lie's theory of continuous groups since P (or Θ in Benedikt's notation) is the matrix (or operator) of an infinitesimal transformation of the rotation group. Cf. L. P. Eisenhart, *Continuous Groups of Transformations* (Princeton, 1933), p. 43. (b) Cf. F. Hausdorff, "Die symbolische Exponentialformel der Gruppentheorie," Leipzig Ber. **58**, 19–48 (1906); and H. Schwerdtfeger, "Sur une formule de H. Poincaré," L'enseignement mathématique **32**, 304–319 (1933), (Sections 5, 6).

According to the laws of matrix calculus the infinite series in (2.4) can be replaced by a finite expression. In fact, by Cayley's identity (1.4), *viz.*, $P^3=-\omega^2P$ all positive powers of P can be expressed by means of P and P^2 only; thus, after an easy calculation, we obtain[6]

$$A=E+(\sin\omega/\omega)P+[(1-\cos\omega)/\omega^2]P^2 \quad (2.5)$$

where $\omega=\theta$ measures the angle of the rotation A in the plane normal to its axis which is defined by the vector p.

3. EULER'S RATIONAL PARAMETRIZATION

The formula (2.5) shows that the elements of the matrix A are non-rational, even transcendental, functions of the canonical parameters p_ν. For many reasons it is desirable to have a set of new parameters (to be introduced by certain functions of the p_ν) such that the elements of A are *rational* functions of the new parameters. This "rationalization" of the representation of the rotation group can be effected in the following way.

The p_1, p_2, p_3 being independent variables it is evident that the group **G** of all rotations constitutes a certain three-dimensional manifold. We shall establish a one-two correspondence between **G** and another three-dimensional manifold which may be described as the unit-sphere **S** in the Euclidean space of four dimensions. Let

$$\mathfrak{u}=\begin{pmatrix}u_0\\u_1\\u_2\\u_3\end{pmatrix}=\begin{pmatrix}u_0\\u\end{pmatrix},\quad \left(u=\begin{pmatrix}u_1\\u_2\\u_3\end{pmatrix}\right), \quad (3.1)$$

be a point in this space. This will be a point on **S** if it satisfies the condition

$$u_0^2+u_1^2+u_2^2+u_3^2=u_0^2+u'u=1 \text{ (or } \mathfrak{u}'\mathfrak{u}=1). \quad (3.2)$$

Subjected to this relation, the u_ν will be a set of rational parameters. We put

$$u_0=\cos\frac{\omega}{2},\quad u=\frac{\sin\omega/2}{\omega}p. \quad (3.3)$$

Hence setting again $U=[u]$, we have

$$P=(\omega/(\sin\omega/2))U,$$

and by means of elementary trigonometrical formulas we find from (2.5)

$$A=E+2u_0U+2U^2. \quad (3.4)$$

[6] For another elementary derivation of this formula see F. D. Murnaghan, reference 1, p. 236–237.

Thus the elements of A are quadratic polynomials in the u_0, u_1, u_2, u_3, the so-called "quaternion parameters" of the rotation group **G**. Since they are four, the relation (3.2) is essential; using it and expressing the matrix U^2 according to (1.5) we can write (3.4) in the form

$$A=(u_0^2-u'u)E+2(u_0U+uu'). \quad (3.5)$$

Hence we see that for every point $\mathfrak{u}$ of the sphere **S** there is exactly one properly orthogonal matrix A. This relation, however, between the points $\mathfrak{u}$ of **S** and the matrices A of the group **G** is certainly not a one-one correspondence since diametrally opposite points $\mathfrak{u}$ and $-\mathfrak{u}$ on **S** give the same A. To show that they are the only ones leading to the same A, let us assume that $\mathfrak{v}$ is a point of **S** that gives the same matrix A as $\mathfrak{u}$ does, such that

$$U^2+u_0U=V^2+v_0V. \quad (3.6)$$

Hence follows

$$(U^2+u_0U)v=0 \quad \text{or} \quad (U+u_0E)Uv=0,$$

or since U is skew-symmetric and therefore $U+u_0E$ regular if $u_0\neq 0$, one has $Uv=0$, i.e., $v=\lambda u$ with a certain real number λ for which, by (3.6), it follows

$$(1-\lambda^2)U^2+(u_0-\lambda v_0)U=0.$$

Hence if we further assume $u_0\neq\pm1$ so that $u\neq 0$, we conclude $\lambda^2=1$, i.e., $v=\pm u$, $v_0=\pm u_0$, or $\mathfrak{v}=\pm\mathfrak{u}$. If $u_0=\pm1$, this is obvious, since then $u=0$. In the limit case $u_0=0$ one has necessarily $u\neq 0$ and from (3.6) follows $v_0=0$ as otherwise the matrix $U^2-V^2=v_0V$ would be symmetric and skew-symmetric. Hence $U^2=V^2$, or $U^2v=0$. As for any non-zero skew-symmetric matrix U also the matrix U^2 is of rank 2, we have again $v=\lambda u$ as only solution and accordingly $\lambda^2=1$. This proves the above statement.

Geometrically the transition from $\mathfrak{u}$ to $-\mathfrak{u}$ means that we replace p by $-p$ and at the same time ω by $2\pi-\omega$; both rotations produced in this way are virtually identical while the corresponding $\mathfrak{u}$'s are opposite in sign.

4. THE COMPOSITION LAW IN QUATERNION PARAMETERS

Let

$$A=E+2u_0U+2U^2,\quad B=E+2v_0V+2V^2$$

be two orthogonal matrices, according to (3.4) represented by means of their quaternion parameters u_ν, v_ν, respectively. To find the com-

position law in these parameters we notice that, apart from its sign, the vector u is completely determined by A. The vector satisfies the eigenvalue equation

$$Au=u, \tag{4.1}$$

which expresses that u coincides with the axis of the rotation and thus is fixed apart from a factor ρ the square of which can be found from (3.2) if we introduce the value of u_0 given by (3.3). The same can be seen from the fact that $2u_0U$ is the (uniquely determined) skew-symmetric component of the matrix A, i.e., $\frac{1}{2}(A-A')$.

Now let us consider the rotation matrix

$$BA=E+2w_0W+2W^2, \quad W=[w].$$

One has to express w_0 and w by means of u_0, u, v_0, v. To do this without circumstantial calculation we state first that

$$w^*=u_0v+v_0u+Vu \tag{4.2}$$

is the axis vector of the matrix BA and therefore

$$w=\mu w^*,$$

where μ is a real factor still unknown. In fact, with respect to (4.1) and making use of the simple formulas of Section 1 we find easily:

$$\begin{aligned} Aw^* &= u_0v+2u_0^2Uv+2u_0U^2v+Vu \\ &\quad +v_0u+2U^2Vu+2u_0UVu \\ &= u_0v+v_0u+Uv, \end{aligned}$$

and thus by means of symmetry

$$BAw^*=u_0v+v_0u+Vu=w^*.$$

To determine the unknown factor μ we use the condition $w_0^2+w'w=1$ (cf. (3.2)); hence

$$\begin{aligned} w_0^2 &= 1-\mu^2w^{*\prime}w^* \\ &= u_0^2+u'u-\mu^2(u_0^2v'v+v_0^2u'u+2u_0v_0\cdot u'v \\ &\qquad +u'u\cdot v'v-(u'v)^2) \\ &= u_0^2(1-\mu^2\cdot v'v)+u'u(1-\mu^2(v_0^2+v'v)) \\ &\qquad -2\mu^2u_0v_0\cdot u'v+\mu^2(u'v)^2 \\ &= 1-\mu^2+\mu^2(u_0v_0-u'v)^2 \end{aligned}$$

which, evidently, represents w_0 as a rational function of the u, and of the v, if, and only if, $\mu^2=1$. Thus $w=\pm w^*$. Therefore the composition law in the quaternion parameters is

$$\left.\begin{aligned} w_0 &= u_0v_0-u'v \\ w &= u_0v+v_0u+Vu. \end{aligned}\right\} \tag{4.3}$$

The four-coordinate columns

$$\mathfrak{u}=\begin{pmatrix} u_0 \\ u \end{pmatrix}, \quad \mathfrak{v}=\begin{pmatrix} v_0 \\ v \end{pmatrix}, \quad \mathfrak{w}=\begin{pmatrix} w_0 \\ w \end{pmatrix}$$

may be considered as representatives of real quaternions. It is readily verified that (4.3) then is, by definition, the interpretation of the quaternion equation

$$\mathfrak{w}=\mathfrak{v}\circ\mathfrak{u} \tag{4.4}$$

if the $\circ$ indicates the quaternion multiplication. In this way we recognize a two-one isomorphism between the group **Q** of all real quaternions $\mathfrak{u}$ of norm one (i.e., $\mathfrak{u}'\mathfrak{u}=u_0^2+u'u=1$) (cf. (3.2)) and the rotation group **G**. One also says that **Q** is a double-valued representation of **G**.

5. THE UNITARY REPRESENTATION

From (4.3) we derive at once another representation of **G**. By means of the real quaternion parameters u, we form the two-rowed matrix with complex elements

$$T=T(\mathfrak{u})=\begin{pmatrix} u_0+iu_2 & -u_1+iu_3 \\ u_1+iu_3 & u_0-iu_2 \end{pmatrix}. \tag{5.1}$$

Because of the condition (3.2) its determinant is

$$|T|=u_0^2+u'u=1. \tag{5.2}$$

Further the matrix T is unitary: Denoting by $\bar{T}$ the matrix with conjugate complex elements, one has $\bar{T}'T=E_2$ (i.e., the two-rowed unit matrix). All these matrices $T(\mathfrak{u})$ form a group, the so-called "special unitary group" **U**. There is a one-one correspondence between these matrices $T(\mathfrak{u})$ and the real quaternions $\mathfrak{u}$ of norm one: $\mathfrak{u}\leftrightarrow T(\mathfrak{u})$. Moreover this correspondence is at once an isomorphism between the groups **U** and **Q**; in fact the relations (4.3) show that

$$T(\mathfrak{v})T(\mathfrak{u})=T(\mathfrak{v}\circ\mathfrak{u}). \tag{5.3}$$

Therefore also **U** is a double-valued representation of the rotation group **G**. This is the fundamental fact of the so-called theory of spinors.[7]

6. THE CAYLEY-KLEIN PARAMETERS

It is of interest to establish the direct relation which exists between the two groups **G** and **U** without the mediation of the quaternion group

[7] A double-valued representation of **G** is also called a spin-representation. For further details on this subject see Chapter X of Murnaghan's book (reference 1); further, B. L. van der Waerden, *Die gruppentheoretische Methode in der Quantenmechanik* (Berlin 1932), Section 16; and E. Cartan, *Leçons sur la théorie des spineurs I, Actualités scientifiques* (Paris, 1938), Chapters III and IV, p. 643. Cf. also E. T. Whittaker, *Analytical Dynamics* (Cambridge, 1937), pp. 8–13.

Q. This can be done by means of the following general principle to be explained here in a special form adapted to the present purpose.

Let $\Xi=\begin{pmatrix}\xi_1 & \xi_2\\ \xi_2 & \xi_3\end{pmatrix}$ be a symmetric matrix with variable complex elements ξ_1, ξ_2, ξ_3. We may identify this matrix with the column $\xi=\begin{pmatrix}\xi_1\\ \xi_2\\ \xi_3\end{pmatrix}$ representing a point in the three-dimensional complex space $\mathbf{C}_3$. Further let $\mathbf{T}$ be a group of two-rowed matrices with complex elements:

$$T=\begin{pmatrix}\delta & -\beta\\ -\gamma & \alpha\end{pmatrix} \quad \text{and} \quad |T|=\alpha\delta-\beta\gamma=1.$$

Then $T^{-1}=\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}$. By the coordinate transformation $a^*=Ta$ the quadratic form $a'\Xi a$ (where a is the two-coordinate column of its variables) is carried in the form which has the matrix

$$\mathrm{H}=T^{1-1}\Xi T^{-1}. \tag{6.1}$$

This equation represents a linear transformation in the space $\mathbf{C}_3$ that may be represented in the form $\eta=D(T)\xi$ where

$$D(T)=\begin{pmatrix}\alpha^2 & 2\alpha\gamma & \gamma^2\\ \alpha\beta & \alpha\delta+\beta\gamma & \gamma\delta\\ \beta^2 & 2\beta\delta & \delta^2\end{pmatrix}. \tag{6.2}$$

These matrices $D(T)$ form a group $\mathbf{T}^*$ of three-rowed matrices representing the group $\mathbf{T}$. Evidently the matrix $D(T)$ is representative of T as well as of $-T$; thus the representation $\mathbf{T}^*$ of $\mathbf{T}$ is certainly not isomorphic if $-E_2$ is an element of the group $\mathbf{T}$. Since, however, from

$$S'\mathrm{H}S=T'\mathrm{H}T \quad \text{for all} \quad \mathrm{H}=\begin{pmatrix}\eta_1 & \eta_2\\ \eta_2 & \eta_3\end{pmatrix},$$

it follows $S=\pm T$, we conclude that T and $-T$ are the only elements of $\mathbf{T}$ represented by $D(T)$. Therefore:

The group $\mathbf{T}^*$ is an isomorphic representation of $\mathbf{T}$ if and only if $-E_2$ is not an element of $\mathbf{T}$. If $\mathbf{T}$ contains $-E_2$, the correspondence $T\to D(T)$ is a two-one homomorphism or: $\mathbf{T}$ is a double-valued representation of $\mathbf{T}^*$.

(This is likewise seen to be true if $\mathbf{T}$ is a group of n-rowed matrices.)

Because of $|T|=1$ it follows from (6.1) that the determinant of the matrix Ξ, i.e., the quadratic form

$$|\Xi|=\xi_1\xi_3-\xi_2^2=\xi'Q\xi, \tag{6.3}$$

where

$$Q=\begin{pmatrix}0 & 0 & \frac{1}{2}\\ 0 & -1 & 0\\ \frac{1}{2} & 0 & 0\end{pmatrix}$$

is an invariant under all the transformations $\eta=D(T)\xi$.

Now we choose instead of the general group $\mathbf{T}$ the special unitary group $\mathbf{U}$. This means that $\gamma=-\bar\beta$ and $\delta=\bar\alpha$ and therefore

$$T=T(\mathfrak{u})=\begin{pmatrix}\bar\alpha & -\beta\\ \bar\beta & \alpha\end{pmatrix},$$

where, with regard to (5.1),

$$\alpha=u_0-iu_2, \quad \beta=u_1-iu_3, \quad (\bar\alpha\alpha+\bar\beta\beta=1); \tag{6.4}$$

these two complex numbers are the Cayley-Klein parameters. According to (6.2) we have

$$D(T)=\begin{pmatrix}\alpha^2 & -2\alpha\bar\beta & \bar\beta^2\\ \alpha\beta & \alpha\bar\alpha-\beta\bar\beta & -\bar\alpha\bar\beta\\ \beta^2 & 2\bar\alpha\beta & \bar\alpha^2\end{pmatrix}.$$

By applying the non-real substitution

$$\xi_1=x_1+ix_3, \quad \xi_2=-ix_2, \quad \xi_3=x_1-ix_3, \tag{6.5}$$

the invariant form $|\Xi|$ of (6.3) is turned into the "unit form:"

$$|\Xi|=x_1^2+x_2^2+x_3^2=x'x.$$

Hence when denoting by R the matrix of the substitution (6.5), i.e.,

$$R=\begin{pmatrix}1 & 0 & i\\ 0 & -i & 0\\ 1 & 0 & -i\end{pmatrix},$$

the matrices $R^{-1}D(T)R$ leave invariant the form $x'x$, and therefore are orthogonal. Explicit calculation yields

$$R^{-1}D(T)R=\begin{pmatrix}\frac{1}{2}(\alpha^2+\bar\alpha^2+\beta^2+\bar\beta^2) & i(\alpha\beta-\bar\alpha\bar\beta) & \frac{i}{2}(\alpha^2-\bar\alpha^2+\beta^2-\bar\beta^2)\\ i(\alpha\beta-\bar\alpha\bar\beta) & \alpha\bar\alpha-\beta\bar\beta & -(\alpha\beta+\bar\alpha\bar\beta)\\ -\frac{i}{2}(\alpha^2-\bar\alpha^2-\beta^2+\bar\beta^2) & \alpha\bar\beta+\bar\alpha\beta & \frac{1}{2}(\alpha^2+\bar\alpha^2-\beta^2-\bar\beta^2)\end{pmatrix}.$$

which is first seen to be real and, moreover, because of (6.4), coincides with the matrix A as given by (3.5). Since $\mathbf{U}$ contains the matrix $-E_2$ it follows again that $\mathbf{U}$ is a double-valued representation of the rotation group $\mathbf{G}$.

The composition of the Cayley-Klein parameters follows from the multiplication law of the matrices $T(\mathfrak{u})$. If α_1, β_1 are the parameters of $T(\mathfrak{v})$, those of the product $T(\mathfrak{v})\cdot T(\mathfrak{u})$ are

$$\alpha_2=\alpha_1\alpha-\bar{\beta}_1\beta, \quad \beta_2=\bar{\alpha}_1\beta+\beta_1\alpha. \tag{6.6}$$

7. CAYLEY'S RATIONAL PARAMETRIZATION

There is another rational representation of the general proper orthogonal matrix A where likewise a general skew-symmetric matrix is involved: If Q is skew-symmetric then

$$A=(E-Q)(E+Q)^{-1}=(E+Q)^{-1}(E-Q) \tag{7.1}$$

is proper orthogonal. In fact, as

$$|E+Q|=|E+Q'|=|E-Q|$$

is always different from zero, one has $|A|=+1$. Further $A'=(E+Q)(E-Q)^{-1}$ and therefore $A'A=E$.

Conversely, if A is any matrix such that

$$|E+A|\neq 0, \tag{7.2}$$

one can solve the Eq. (7.1) with respect to Q; it is

$$Q=(E-A)(E+A)^{-1}=(E+A)^{-1}(E-A). \tag{7.3}$$

If, moreover, A is orthogonal, the matrix Q given by (7.3) is skew-symmetric. To prove this we write the second relation (7.3) in the form $(E+A)Q=E-A$; by transposition we get $Q'(E+A')=E-A'$ whence by multiplication by A we obtain

$$Q'(E+A)=-(E-A).$$

This proves the statement.

Thus, for every skew-symmetric Q we get a proper orthogonal matrix A, while for those, and only those, proper orthogonal matrices A that satisfy the condition (7.2), a skew-symmetric Q can be found such that (7.1) holds true. Therefore Cayley's parametrization (7.1) of the orthogonal group is incomplete: It omits the matrices A for which $|E+A|=0$.[8]

So far this consideration is valid for the orthogonal group in any number n of dimensions. In the case $n=3$ we shall show that the natural completion of Cayley's parametrization is provided by Euler's representation (3.4). From (7.1) we have

$$A(E+Q)=E-Q; \tag{7.4}$$

thus, when denoting by q the vector column with which the skew-symmetric matrix Q is associated ($Q=[q]$), we have $Aq=q$ which means that q coincides with the axis of the rotation A. Hence we have

$$q=\lambda u$$

and (3.4) and (7.4) give the following condition for λ:

$$(E+2u_0U+2U^2)(E+\lambda U)=E-\lambda U,$$

which by means of (1.4) is reduced to

$$2(u_0+\lambda-\lambda\cdot u'u)U+2(1+\lambda u_0)U^2=0.$$

As the symmetric and the skew-symmetric component of the zero-matrix 0 both are 0, we conclude

$$1+\lambda u_0=0,$$

except when $U=0$. Hence λ can be determined if, and only if, $u_0\neq 0$, i.e., ω, the rotation angle of A, is not an odd multiple of π (cf. (3.3)).

---

[8] For a discussion of the question of the "exceptional elements" in Cayley's parametrization for the rotation group in n dimensions see H. Weyl, *The Classical Groups* Chapter II, B. Further the paper by A. Danilewsky, "Sur une generalisation des formules de Cayley" (Ukrainian, French summary), Commun. Inst. Sci. Math. et Mécan. Univ. Charkow et Soc. Math. Charkow, (IV) **14**, 85–94 (1937). Reviewed in Zentralblatt für Mathematik **19**, 244 (1939). For further references see the article by A. Loewy, "Kombinatorik, Determinanten und Matrices," in E. Pascal's *Repertorium der höheren Mathematik* (Leipzig-Berlin, 1910), second edition, Chap. II, p. 133; and C. C. MacDuffee, *Theory of Matrices, Ergebnisse der Mathematik* (Berlin, 1933), Vol. II, 5, p. 79, note 5. For $n=4$ and 5, cf. the German edition of M. Bocher's *Introduction to Higher Algebra, viz., Einführung in die höhere Algebra* (Leipzig-Berlin, 1925), p. 340.

Here and There

New Appointments

Lt. Kenneth W. Moore, formerly chief of the photographic section of the Armour Research Foundation, Chicago, has joined the staff of the Midwest Research Institute, Kansas City, and will be in charge of the photography and optics department.

The Sharples Corporation of Philadelphia has added the following physicists to the staff of its Research Laboratories: Malcolm H. Hebb, Roderic M. Scott, and Robert E. Payne. All three have been at the Underwater Sound Laboratory of Harvard University.

Miriam Lauren, formerly of the Rockefeller Institute for Medical Research, has joined the staff of Foster D. Snell, Inc., where she will have charge of a greatly enlarged microanalytical laboratory. Gerald M. Compeau, recently research analytical chemist with the Colgate-Palmolive Peet Company, has also joined the staff of Foster D. Snell, as assistant to the Director of the Analytical Laboratory.

Dr. Carleton C. Murdock, professor of physics at Cornell University, has been elected dean of the faculty. He succeeds Dr. Cornelius Betten, who recently retired.

William I. Book, professor of physics at the University of Pennsylvania, has been appointed professor emeritus. He retired from active duty on July 1.

Dr. George H. Burnham, associate editor of *The Review of Sciencific Instruments* and for the last three years assistant to the director of the American Institute of Physics, has been appointed associate professor of physics and assistant to the president of Norwich University, Northfield, Vermont.

Joyce C. Stearns, former professor of physics, University of Denver, has been appointed to the newly created post of Dean of the Faculties, Washington University, at St. Louis, Missouri.

N. Frank Stump, formerly of the General Personnel Department, Revere Copper and Brass, Inc., recently joined the Industrial Vision Department of the Bausch and Lomb Optical Company.

Reuben G. Gustavson has resigned as president of the University of Colorado to become vice president of the University of Chicago.

The June *Technology Review* reported the following promotions in the Physics Department at Massachusetts Institute of Technology: Robley D. Evans to Professor and Herman Feshbach, Laszla Tisga, and Arthur F. Kip to Assistant Professors.

Washington State Agriculture Demonstrations

The United States Department of Agriculture's Extension Service, in cooperation with Washington State College, recently completed its second original wartime showing of the results of research. These results were shown through demonstrations of "Guides to Successful Farming," as follows: fruit and vegetable storage; model milk house; refrigeration; moving picture machine; agricultural equipment; sound equipment; utility room; tools and rust control; 4-H Clubs; economic farm unit; bulletins; rat control; hay drier; light metals; horticulture; poultry; food carnival; farm corral; lighting. Approximately 40,000 men and women, as well as many children and teen-age youths, attended 36 different demonstrations throughout the State of Washington.

Engineering Physics News at Ohio State

The June 1945 issue of the *Engineering Experiment Station News* of Ohio State University is devoted to the development of engineering physics as a profession. The papers in this issue are as follows: "Engineering Physics—A New Profession" by Alpheus W. Smith; "The Engineering Physicist in the Industrial Laboratory" by James W. Hackett; "The Renaissance of the 'Gaslight' Era" by Erwin F. Lowry; "Engineering Opportunities in the Instrument Industry" by Ralph E. Clarridge; "The Physicist in the Petroleum Industry" by Frank Morgan; "Some Problems in the Manufacture of Surgical Dressings" by Herbert M. Strong; "Physics in the Rubber Industry" by William C. Sears.

Enemy Patents at Midwest Research Institute

Dexter North, chief of patent administration of the Alien Property Custodian's Office, has turned over to Midwest Research Institute, Kansas City, Missouri, complete working drawings on the Office's huge collection of patents seized from enemy aliens and nationals of occupied countries. Mr. North stated that the government is anxious to "put these enemy patents to work for the benefit of American economy." Already 9,366 of the inventions have been licensed to 760 American firms and individuals. Products valued at more than $150,000,000, largely war goods, have already been produced from them. Rights for lifetime use of patents of enemy countries may be secured for a fee of $15. The same fee is charged by the Alien Property Custodian for use of patents originating in non-enemy and occupied countries, plus a small royalty. The file is accessible at all times to manufacturers, agriculturists, educational institutions, and others.

Suit for Electron Microscopy Patents

Reinhold Rudenberg, formerly a citizen of Germany but now a naturalized citizen of the United States and professor of electrical engineering at Harvard University, has brought suit against the Alien Property Custodian to regain possession of his patents, which are said to be fundamental to electron microscopy.

Meeting of Electron Microscope Society

The Electron Microscope Society of America will hold its third annual meeting November 30 and December 1 at the Frick Chemical Laboratory of Princeton University. The purpose of the meeting is to acquaint the members with the rapid advances in the past year in the applications and techniques of electron microscopy and diffraction in currently important problems in physics, chemistry, and biology. The Program Chairman is John Turkevich, Chemistry Department, Princeton University, Princeton, New Jersey.

Expansion of Polymer Bulletin

Interscience Publishers, Inc., and the Bureau of High Polymer Research at the Polytechnic Institute of Brooklyn, cooperating publishers of the *Polymer Bulletin*, have decided to extend its scope and program and to transform it, on January 1, 1946, into an international *Journal for High Polymer Research*. This journal will be devoted to the publication of original papers and

surveys in the entire field of the chemistry and physics of high polymeric substances. It will aim to reflect mainly the scientific progress and development in the field of plastics and elastics all over the world. The Editor will be Dr. H. Mark, and Assistant Editors will be Drs. Paul M. Doty and Charles C. Price.

Kansas City Technical Societies Council

Seventeen professional and technical societies of the Kansas City area have joined forces in establishing the Technical Societies Council for unified action in matters pertaining to the development of scientific endeavors. A monthly bulletin will be published to program activities of all organizations. Some of the principal objectives to be advanced immediately include support for establishing better technical and scientific libraries in that area; to encourage the establishment of educational courses of a technical and scientific nature in the schools of this area and coordinate the vocational guidance activities of the constituent organizations; to cooperate with civic, educational and government agencies in matters requiring professional assistance. Headquarters of the new Council will be in the offices of the Midwest Research Institute, Kansas City 2, Missouri.

RCA Scholarships

Science reports that a scholarship plan for the encouragement of promising young scientific students has been adopted by the Radio Corporation of America. The plan provides for ten students to receive RCA scholarships during the academic year 1945–1946, thirty during 1946–1947, fifty during 1947–1948, and sixty each academic year thereafter. Each scholarship consists of a cash award of $600. Those eligible will include all students enrolled at universities to be selected by the RCA Education Committee. Selection of students will be made upon recommendation of the dean of the specified university and approval by the committee. The chairman of the Education Committee is Dr. James Rowland Angell, president emeritus of Yale University and public service counselor of the National Broadcasting Company.

Research Men Honored at Harvard

At its commencement exercises Harvard University conferred the degree of doctor of science upon Frederick E. Terman, head of the department of electrical engineering of Stanford University, now on leave as director of radio research for war at Harvard University, and Frederick V. Hunt, professor of physics and director of the Underwater Sound Laboratory of Harvard University.

Physics Instructors Abroad

The Office of Scientific Personnel of the National Research Council has announced that the following men are teaching physics to American soldiers attending the Army school at Shrivenham, England: F. W. Sears, in charge, G. H. Cameron, C. D. Shallenburger, H. L. Lein, R. B. Sherman, Captain W. R. Hazeltine, Lt. M. J. Gottlieb; and at Biarritz, France: Paul Gleason, in charge, R. B. Kennard, Franklin Furst, E. H. Kinsey, and Major Sherman E. Eager.

Joint Board on Scientific Information Policy

According to *Science*, the facilities of the Office of War Information, Domestic Branch, have been made available to the newly organized Joint Board on Scientific Information Policy for the handling of information on scientific and technical developments of the Government during the war. The Board will be responsible for the organization and release of scientific and allied information for the Office of Scientific Research and Development, the War and Navy Departments, and the National Advisory Committee for Aeronautics. Board chairman is Dr. John T. Tate. A committee consisting of the Secretaries of War and the director of the Office of Scientific Research and Development, in recommending the formation of the Board, made this statement: "A positive information policy is desirable to supplement spot news stories, which probably can never be controlled as to accuracy of content or interpretation of scientific development. The basic purpose of such a policy is to establish in the minds of the public a proper concept of how scientific development really has taken place. . . . The work of carrying out such a policy should be entrusted to a small group of personnel whose duties shall include determining suitable stories, getting the necessary background of information from the laboratories and operational units concerned, and in general guiding the story to a release point." The basic reports of the Board will include such fundamental subjects as radar rockets, metallurgy, jet propulsion, aircraft design, public health, medicine, plastics.

Audar, Inc.

John Meck, president of the John Meck Industries, Inc., of Plymouth, Indiana, recently announced the formation of Audar, Inc., a separate corporation which will manufacture and sell public address systems and audio amplifiers as an affiliate of John Meck Industries, Inc. The officers of the corporation are John S. Meck, president, E. W. Applebaum, treasurer and general manager, and Russell G. Eggo, secretary.

Awarded Honorary Degree

Captain Ralph D. Bennett, Technical Director of Naval Ordnance Laboratory, Navy Yard, Washington, D. C., was awarded an honorary Doctor of Science degree at Union College, his alma mater, on June 24. Before entering the service of his country, Captain Bennett was Professor of Electrical Measurements at the Massachusetts Institute of Technology.

Plans for Chinese Institute

Equipment for a National Institute of Physiology in China, which will be unique among scientific institutions in the world, is now being purchased in the United States by its future director, Dr. S. C. Shen, who recently came to the Polytechnic Institute of Brooklyn to participate in a new project which takes the tools of biochemistry into the high-polymer field. Dr. Shen is in this country as a representative of the Ministry of Education of the Chinese government.

It is planned to locate the new institute in Shanghai, since it is considered advisable to have it near a seaport. According to Dr. Shen, who will take China's first ultracentrifuge back with him and who is now purchasing basic equipment which is being shipped out as rapidly as possible, China will break from tradition in establishing a National Institute devoted entirely to the study of developmental physiology, a study which always has been part of the curriculum of the medical school in this country and Europe. The first project of the Institute will be an investigation of the nutritional status of the Chinese with a view to improving the content of the foods with essential minerals and vitamins according to procedures already current in this country.

Navy Research Laboratories at Pennsylvania State College

The U. S. Navy has completed arrangements for the establishment of two permanent ordnance research laboratories which will be in operation by fall at The Pennsylvania State College under a Bureau of Ordnance contract. The establishment of these laboratories is in keeping with the Bureau of Ordnance policy of forming research affilia-

tions with educational institutions to further development of naval weapons and for educational purposes in such military problems during the postwar period.

The Ordnance Research Laboratory, dealing with underwater ordnance, will continue the research formerly carried on by the Underwater Sound Laboratory at Harvard University which has been sponsored by the Office of Scientific Research and Development, with technical direction from the Bureau of Ordnance. Dr. Eric A. Walker, who was in charge of the Ordnance Research Division at Harvard, will direct the new laboratory and also head the Electrical Engineering Department of the College. A staff of approximately 125 scientists, technicians, and clerks will be transferred from the Harvard Laboratory to the new unit.

The Petroleum Refining Laboratory, which has been in operation for sixteen years at The Pennsylvania State College, will continue with petroleum research under a direct contract with the Bureau of Ordnance. Dr. M. R. Fenske will be director of the Laboratory, a position which he has held since it was established.

The Ordnance Research Laboratory has been placed under the School of Engineering, and the Petroleum Refining Laboratory is under the School of Chemistry and Physics. Most of the scientific and technical personnel will hold academic appointments on the faculty of the College.

Gerard Swope Scholarship Fund

To assist in higher education and fundamental research work in scientific and industrial fields, a fund of $400,000 has been set aside by the directors of the General Electric Company to be known as the Gerard Swope Foundation, in honor of Mr. Swope, who served as president of General Electric for more than 19 years and as a director of the company for more than 22 years. It is proposed to utilize the income of the Foundation in the following ways:

In granting loans or scholarships to employees and to the children of present or former employees of the company and its affiliated companies, to help them pursue their work in any field of study and in any approved university, college, or technical school in the United States that they may select; in granting graduate fellowships to the same classifications of individuals to be used by the recipient for graduate work in industrial management, engineering, the physical sciences, and in any other scientific or industrial field; in granting loans, scholarships, or fellowships in any of the above fields to any other person who may be deemed worthy of assistance; and finally, in granting graduate fellowships, to make contributions to the university, college, or technical school to help in defraying the cost of equipment and material, especially needed in connection with a research project.

Applicants for loans, scholarships, and fellowships will be considered not only from the standpoint of academic attainment and financial needs, but also as to character and personal qualifications for deriving the greatest good from the continuation of their studies. The grants will be made without regard to sex, color, creed, or national origin. It is planned to make a statement of awards under the foundation in each of the company's annual reports.

RCA Tube Division

Organization of the RCA Tube Division, to embrace all electron tube activities of the company, with L. W. Teegarden as general manager, has been announced by Frank M. Folsom, executive vice president of RCA in charge of RCA Victor Division. The new Tube Division provides for complete integration and coordination of RCA's tube and tube equipment engineering, manufacturing, and sales activities. It is planned to transfer the division's headquarters from Camden to Harrison, New Jersey, in the early future.

Bulletin of Mathematical Biophysics

The table of contents of the September 1945 *Bulletin of Mathematical Biophysics* contains the following:

The Physiological Factors Which Govern Inert Gas Exchange—M. F. MORALES AND R. E. SMITH.
Contribution to the Theory of Discrimination Learning—T. L. HILL AND L. E. HILL.
The Mathematical Biophysics of Some Mental Phenomena—N. RASHEVSKY.
Mathematical Biophysics of Abstraction and Logical Thinking—N. RASHEVSKY.
The Differential Difference Equation for Epidemics—J. ERNEST WILKINS, JR.
A Reinterpretation of the Mathematical Biophysics of the Central Nervous System in the Light of Neurophysiological Findings—N. RASHEVSKY.
University of Chicago Press, Chicago, Illinois, Volume 7, Number 3.

Products of Tomorrow Exposition

America's first annual "Products of Tomorrow Exposition" is being planned for an opening in Chicago on January 18, 1946, subject to military exigencies. It will be located in the Coliseum, Armory, and other exhibition halls already in existence and ready for use. The "Exposition" will be an annual event to provide a springboard for the introduction of new products and new designs by all industry. It will be a private enterprise in itself, with no dependence upon municipal or government facilities and personnel.

R. T. P. Translations

The Durand Reprinting Committee, a non-profit organization set up by several members of the Guggenheim Aeronautical Laboratory at the California Institute of Technology, has been organized at Pasadena for the purpose of reprinting and distributing, on a yearly subscription basis, the R. T. P. translations of the British Ministry of Aircraft Production. Composed of translations of articles appearing in both enemy and allied publications, the papers will be of interest to aircraft companies, engine manufacturers, research laboratories, colleges, etc. Quotations on the reprints will be furnished upon request. Following are the titles of the first group of reprints released by the Committee:

The Distribution of Forces on the Rivet in Stretched Joints with Constant Strap Cross Section.
The Theory of the Combustion Process.
Recent Development of the Two-Stroke Engine, II, Design Features.
Torque Meter for Power-Driven Aircraft Models in the Wind Tunnel.
Experiments on the Influence of Air Content on Cavitation and Corrosion, in the Case of Water Flow.
Harmonic Analysis of Discrete Numerical Series.
Notes on a Controlled Type of Unidimensional Motion.
The Calculation of the Lift Distribution of Aerofoils.
Discharge Tests and Force Measurements on the Stationary Blading of a Single-Stage Experimental Turbine in the Air Test Plant.
Notch Effect.
Modern Plastics as Raw Materials for Adhesives.
The Kutzbach Method for the Graphical Analysis of Natural Torsional Oscillations.

Research Conference on X-Ray and Electron Diffraction

On Friday and Saturday, November 2 and 3, 1945, a research conference on X-Ray and Electron Diffraction will be held by the University of Pittsburgh, Pittsburgh, Pennsylvania at the Mellon Institute Auditorium. All interested persons are welcome to attend. The program is as follows:

Friday, November 2, 2:00 P.M.

1. "X-ray Diffraction Studies of the System Zn_2TiO_4—$NiTiO_3$. Zinc Orthotitanate and Nickel Metatitanate." HERMANN BIRNBAUM AND R. K. SCOTT, *Cooperative X-Ray Laboratory, University of Pittsburgh.*
2. "X-Ray Diffraction Studies of Synthetic Triacid Triglycerides." L. J. FILER, JR., S. S. SIDHU, CHIADAO CHEN, AND B. F. DAUBERT, *University of Pittsburgh.*
3. "X-Ray Diffraction Analysis of Fisher-Tropsch Catalysts." L. J. E. HOFER, *U. S. Bureau of Mines, Pittsburgh, Pennsylvania.*

4. "Device and Method for Quantitative Analysis by X-Ray Diffraction." STANLEY A. BROSKY, *Radiographic Laboratory, Pittsburgh Testing Laboratories.*
5. "X-Ray Diffraction Studies of Alkali Metal Iron Sulfates—the Systems $K_3Fe(SO_4)_3-Na_3Fe(SO_4)_3$ and $KFe(SO_4)_2-NaFe(SO_4)_2$." S. S. SIDHU, *University of Pittsburgh* AND RICHARD C. COREY, *Research and Development Department, Combustion Engineering Company*, New York, New York.

6:00 P.M.

Testimonial Dinner in honor of Professor Wheeler P. Davey of Pennsylvania State College will be held at the Hotel Webster Hall. Persons desiring to attend should make reservations with Dr. S. S. Sidhu, University of Pittsburgh, before the first of November.

8:30 P.M.

1. "The Use of Fournier Syntheses in Crystal Structure Analysis." MAURICE L. HUGGINS, *Research Laboratories, Eastman Kodak Company, Rochester, New York.*

Saturday, November 3, 9:30 A.M.

1. "Microscopy by Using X-Ray Diffraction." C. S. BARRETT, *Carnegie Institute of Technology.*
2. "New Techniques in Electron Microscopy as Applied to Sectioning of Solid Materials and Special Methods of Dispersion of Pigments." HAROLD C. O'BRIEN, JR., *University of Pittsburgh.*
3. "The Electron Diffraction Method." E. A. GULBRANSON, *Research Laboratories, Westinghouse Electric Corporation, East Pittsburgh, Pennsylvania.*
4. "Electron Microscope and Electron Diffraction Studies of Oxide Films Formed on Metals and Alloys." R. T. PHELPS, *Research Laboratories, Westinghouse Electric Corporation, East Pittsburgh, Pennsylvania.*
5. "Electron Diffraction Study of Oxide Films Formed on Metals and Alloys at Elevated Temperatures." J. W. HICKMAN, *Research Laboratories, Westinghouse Electric Corporation, East Pittsburgh, Pennsylvania.*

Saturday Afternoon—Visits to Laboratories

Symposium on Industrial Radiography and X-Ray Diffraction

A symposium on Industrial Radiography and X-Ray Diffraction will take place at Marquette University, Milwaukee, Wisconsin, November 9 and 10, 1945. Following is the program:

Friday, November 9, 1945

Afternoon Session

Chairman, FR. J. F. CARROLL,
Head of Physics Department, Marquette University.

Address of Welcome: FR. PETER BROOKS, *President, Marquette University.*
Main Address: DR. OTTO GLASSER, *Cleveland Clinic Foundation, Cleveland*, "Half a Century of Roentgen Rays."

Evening Session

Chairman, FR. J. F. CARROLL

Address: DR. G. L. CLARK, *University of Illinois, Urbana, Illinois*

Saturday, November 10

(Double sessions throughout the day)

Morning Sessions

9:30 MR. S. E. MADDIGAN, *Chase Brass & Copper Company, Waterbury, Connecticut*, "Microradiography."
DR. G. D. ADAMS, *University of Illinois, Urbana, Illinois*, "Use of the Betatron in Radiography."
10:45 MR. J. J. CHYLE, *A. O. Smith Corporation, Milwaukee, Wisconsin*, "Radiographic Quality Control of Metallic Arc Welds."
MR. E. R. THILO, *Laboratory Division, Frankford Arsenal, Philadelphia*, "High Speed Radiography."

Afternoon Sessions

1:30 MR. M. B. EVANS, *X-Ray Incorporated, Detroit, Michigan*, "Determination of X-Ray Technique."
DR. C. S. BARRETT, *Carnegie Institute of Technology, Pittsburgh, Pennsylvania*, "Some Questions That Can Be Answered by X-Ray Diffraction."
2:45 DR. KENT R. VAN HORN, *Aluminum Company of America, Cleveland, Ohio*, "The Identification and Significance of the Discontinuities That May Occur in Aluminum and Magnesium Alloy Castings."
DR. J. N. MRGUDICH, *Burgess Battery Company, Freeport, Illinois*, "Industrial Applications of Spectroscopy, X-Ray Diffraction, and Electron Microscopy."
4:00 MR. DON MCCUTCHEON, *Ford Motor Company, Dearborn, Michigan*, "Radiographic Stress Analysis."
MR. F. G. FIRTH, *North American Philips, Dobbs Ferry, New York*, "The Geiger Counter X-Ray Spectrometer."

Evening Session

Banquet: Speaker: DR. HENRY T. HEALD, *President*, Illinois Institute of Technology, Chicago, Illinois.

Predoctoral Fellowships in the Natural Sciences

The National Research Council announces that it is now ready to receive nominations and applications for the predoctoral fellowships in the natural (i.e., mathematical, physical, and biological) sciences which it is administering under a grant from the Rockefeller Foundation. These fellowships are intended to assist young men and women, whose graduate study has been prevented or interrupted by the war, to complete their work for the doctorate. It is hoped that these fellowships will do much to accelerate the recovery of the scientific vigor and competence of the country which is so seriously threatened by the loss of almost two graduate school generations of scientifically trained men and women.

This program will be administered by a Committee on Predoctoral Fellowships of the National Research Council whose members are Henry A. Barton, Charles W. Bray, Detlev W. Bronk, Luther P. Eisenhart, Ross G. Harrison (Chairman—National Research Council, *ex officio*), W. A. Noyes, Jr., and John T. Tate, chairman: Enid Hannaford, secretary.

The annual stipend will be $1200 for single persons and $1800 for married men. In general it is expected that each recipient will spend at least eleven months per year on academic work. An additional allowance up to $500 per year will be made for tuition fees. Fellowships granted to individuals who are eligible for educational support from the "G.I. Bill of Rights" will be at such stipends as to bring the total income from these two sources to that which would be received at the above rates.

Each fellow, before entering on his graduate studies, will submit for review by the Committee on Predoctoral Fellowships a schedule, approved by the dean of his graduate school, for the completion of his work for the doctorate. This schedule, as approved by the committee, will constitute an informal agreement upon the basis of which stipend payments will be made. At the discretion of the university concerned the fellowship stipend may be supplemented by university grants. All such supplementary sources of income should be made a matter of record with the committee. The progress of the fellows will be subject to periodic review by the committee which reserves the right to cancel fellowships when in their judgment satisfactory progress is not being maintained.

Prospective candidates for these fellowships are urged to apply at once even though they may be unable to undertake their graduate study in the immediate future. Information concerning these fellowships and nomination-application blanks are being mailed out widely to graduate schools and wartime research laboratories. They may also be obtained by writing directly to the Secretary, Committee on Predoctoral Fellowships, National Research Council, 2101 Constitution Avenue N.W., Washington 25, D. C.

New Booklets

The Educational Division of Bell and Howell Company, manufacturer of still and motion picture cameras, projectors, and equipment, has published a booklet entitled *Architects' Visual Equipment Handbook* which is designed to show architects, school boards, church organizations, and others how to plan or adapt rooms for showing of

motion pictures or other visual aids for instruction or entertainment. 20 pages. Free on request to Bell and Howell Company, 7100 McCormick Road, Chicago 45, Illinois.

Westinghouse *Newsfront* for July, 1945 contains highly readable articles on a new plastic protective coating called Fosterite, a miniature short-wave radio set that "broadcasts" the ability of a prospective airplane pilot to compensate for changes in altitude, electric fish fences, and the "big aisle" at Westinghouse where the largest electrical units produced by the company are built. 4 pages, illustrated with photographs and cartoons. Address Westinghouse *Newsfront* at 306 Fourth Avenue, Box 1017, Pittsburgh 30, Pennsylvania.

The Lilly Research Laboratories, Indianapolis 6, Indiana, have devoted the Spring 1945 issue of *Research Today* to the subject of penicillin, its discovery, development, and present methods of production in quantity. 40 pages, many colorful illustrations.

Allis-Chalmers Manufacturing Company has published a 20-page booklet entitled *Introduction to Electronics*, written by Dr. Walther Richter. The booklet, E6358, is available upon request from Allis-Chalmers, 568, Milwaukee 1, Wisconsin.

American Optical Company, Scientific Instrument Division, Buffalo, New York, is offering its Bulletin M136, describing the new Spencer Polarizing Microscope, No. 42. The optical system of this instrument contains Polaroid material in place of the usual calcite polarizing prisms. 4 pages.

Leeds and Northrup Company has issued a revised publication describing a basic electrical-measuring instrument, the students' potentiometer. Designed for teaching, it embodies the potentiometer principle in its simplest form, but is also of sufficient accuracy for serious investigational work in the laboratory. Write for Catalog E-50B(1), Leeds and Northrup Company, 4934 Stenton Avenue, Philadelphia 44, Pennsylvania.

Contributed Original Research

A 100-Million Volt Induction Electron Accelerator

W. F. Westendorp and E. E. Charlton
Research Laboratory, General Electric Company, Schenectady, New York
(Received July 9, 1945)

The plan and construction of a device capable of accelerating electrons to energies as high as 100 million electron volts are described together with the special building to house it. A more detailed description of the device and its construction will appear in the *General Electric Review*. The accelerator has a pole face 76 inches in diameter, weighs 130 tons, and operates on 60-cycle current requiring at full load 200 kilowatts. The machine is air cooled and is capable of continuous operation at full voltage. The toroidal glass vacuum tube of oval section has an external diameter of 74 inches and consists of 16 sectors cemented together. At 100 million volts the x-ray output, measured in a thick-walled (½ inch lead) ionization chamber, is 2600 Roentgens per minute at one meter target-thimble distance and the half-value width of the x-ray beam is 2.0 degrees. The penetrating power of the x-rays in iron at various voltages has been measured and the radiographic possibilities studied.

I. INTRODUCTION

SOME of the earliest ideas on the method of accelerating electrons by means of a time-varying magnetic field have been proposed by Wideröe,[1] Walton,[2] Jassinsky,[3] and Steenbeck.[4]

Prior to the publication of D. W. Kerst's first paper[5] on this subject there was in the literature nothing to indicate that a successful device based on this principle had ever been produced.

[1] R. Wideröe, "Über ein neues Prinzip zur Herstellung hoher Spannungen," Archiv f. Elektrotechnik **21**, 387–406 (1928).
[2] E. T. S. Walton, "The Production of High-speed Electrons by Indirect Means," Proc. Camb. Phil. Soc. **25**, Pt IV, 469–481 (October, 1929).
[3] W. W. Jassinsky, "Beschleunigung der Elektronen im elektromagnetische Wechselstromfeld," Archiv f. Elektrotechnik **30**, 590–603 (1936).
[4] M. Steenbeck, U. S. Patent 2,103,303 (December 28, 1937).
[5] D. W. Kerst, "Letter to the Editor: Acceleration of Electrons by Magnetic Induction," Phys. Rev. **58**, 841 (1940).

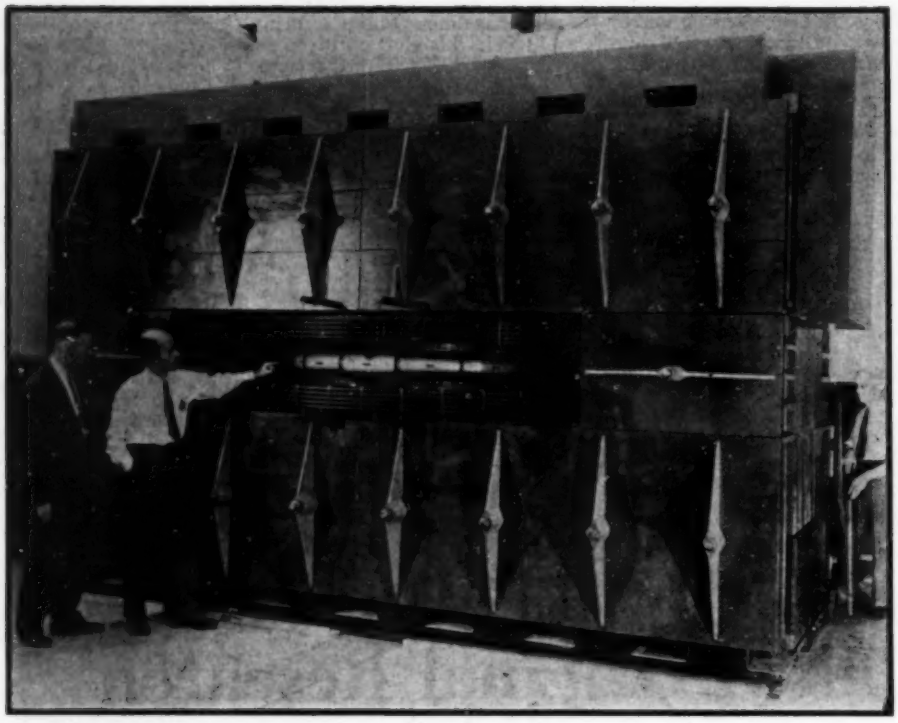


FIG. 1. Front view of electron accelerator.

Since then, a paper has been published by M. Steenbeck[6] describing some of his early experimental work seemingly antedating that of D. W. Kerst.

The general theory of the induction electron accelerator has been described in earlier publications.[4,7–9] The electrons are accelerated by the electric field produced by a time-varying magnetic field. The induction electron accelerator may be compared to the ordinary transformer except that the usual high voltage secondary coil with many turns has been replaced by an evacuated doughnut-shaped glass tube with a hot cathode electron gun. The electrons are shot tangentially into the tube and are caused by the magnetic field to circle around within the tube, acquiring with each revolution the same voltage increase as though they had flowed through a single turn of wire. The electrons are magnetically guided while they are being continuously accelerated in a circular path in the vacuum tube, and at a chosen time they are magnetically guided away from this circular path and made to strike a target and produce x-rays or emerge from the vacuum enclosure as high energy electrons.

Following the development of Dr. Kerst's first successful accelerator, he secured a leave of absence from the University of Illinois and effected an arrangement to cooperate with the staff of the Research Laboratory on the design and construction of a larger one.

The 20-million volt machine resulting from this effort* presented no unexpected difficulties in its construction and encouraged us to proceed with the design and construction of the larger one forming the subject of this paper.

It weighs approximately 130 tons and is about nine feet high, six feet wide, and fifteen feet long. Together with a 24,000 kva capacitor for power factor correction, it is housed in a special building designed to afford protection from the penetrating radiations emitted to occupants of neighboring buildings.

While it is very similar in general design to the earlier 20-million volt machine, its larger size has necessitated many changes both in

[6] M. Steenbeck, "Beschleunigung von Elektronen durch elektrische Wirbelfelder," Naturwiss. **31**, 234–235 (May 7, 1943).

[7] D. W. Kerst, "The Acceleration of Electrons by Magnetic Induction," Phys. Rev. **60**, 47–53 (July, 1941).

[8] D. W. Kerst and R. Serber, "Electronic Orbits in the Induction Accelerator," Phys. Rev. **60**, 53–58 (July, 1941).

[9] D. W. Kerst, "Abstract: Induction Electron Accelerator," Phys. Rev. **59**, 110 (1941).

* This machine immediately following its completion was lent to the University of Illinois.

design and construction. The tube in particular is very different, being made up of 16 molded glass sections cemented end to end.

A front elevation of the accelerator is shown in the photograph of Fig. 1, while a plan view with the upper yoke of the magnet removed, and the lower magnetizing coil and vacuum tube in place, is shown in Fig. 2.

The location of the accelerator in the special building which houses it, is shown in Fig. 3.

Performance

The magnet is operated on 60-cycle current and the flux density at the orbit is 4000 gauss. The electrons are injected with a voltage ranging from 30 to 70 kv and, if allowed to remain in the 66-inch diameter circular orbit for the entire quarter cycle, they circle the magnetic flux about 250,000 times, acquiring on each revolution an average additional energy of about 400 electron volts. As, by pulsing the field, they may be removed from the circular orbit at any time during the quarter cycle, they may be given any desired energy from one million up to the full hundred million electron volts.

As they are made to spiral away from the orbit they impinge upon the target producing x-rays. The device may then be used as a very flexible and easily controlled source of x-rays of any desired voltage up to 100 million. The output, as measured with a thick-walled thimble chamber placed in the center of the x-ray beam, increases rapidly with voltage. Under some conditions it has been as high as 2600 r per minute at 100 million volts, dropping to about seven percent of this value at 20 million.

The half-value width of the x-ray beam is about 12 degrees at 20 million and two degrees at 100 million volts.

Even in the hottest weather the machine may be operated continuously at full load.

FIG. 2. Vacuum tube in position on magnet.

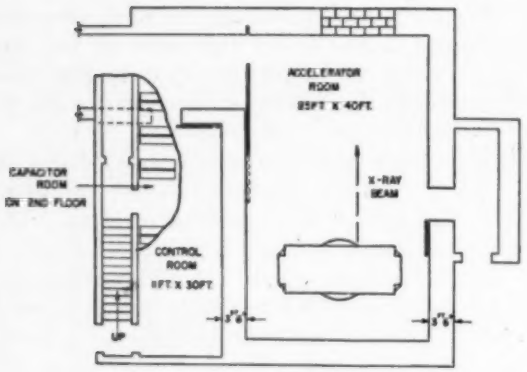


FIG. 3. Floor plan of building housing the electron accelerator.

II. THE MAGNET

The magnet core as shown in Figs. 4 and 5 consists of two yoke pieces spaced apart by two legs, two pole pieces, and two centerpiece disks. These parts are all made of enameled sheet steel (4.5 percent silicon) .014″ thick bonded into slabs or in the case of the pole pieces and centerpiece disks, into pie-shaped sectors. Each slab or sector consists of many sheets stuck together by means of a solventless varnish which is polymerized by baking while the assembly is held in a suitable clamp or form.

The slabs are seven inches thick and are spaced apart, as may be seen in Figs. 2 and 6, by many narrow strips of wood so arranged as to permit of effective forced air cooling. Between hardwood facing planks the slabs are clamped together, as shown in Figs. 1, 2, and 5, by electrically insulated steel bolts.

The Pole Pieces

The sheets of which the pole-piece sectors were built were punched out in the shape of the profile shown in Fig. 7. By subjecting various lengths of pieces in 2″ steps to the same profile die, it was possible to assemble a tapered piece or sector of approximately four degrees in a special clamp. A thin film of solventless varnish was applied to each sheet before assembly and the form bolted together to insure a solid wedge-shaped block. A 15-hour bake at 150°C polymerized the varnish, after which the sectors could be handled as one piece.

The rectangular notches in the ends serve to line up and hold in place the sectors of the bottom pole piece, and to suspend the sectors of the upper pole piece from the top yoke. For this purpose a large split wooden ring, ten in Fig. 8, made up of many pieces of mahogany and accurately machined, is placed around the assembled sectors; similarly, a Textolite ring, six in Fig. 8, is placed in the center in the rectangular notches. A steel spider engages the Textolite ring in the upper pole piece and a brass rod, screwed in the hub of the spider, holds up the center of the pole piece. Steel strips passed between the slabs that make up the upper yoke, engage the wooden ring, and hold up the rim of the pole piece. The lower pole piece is held in place on the bottom yoke by means of a Textolite spider which allows circulation of cooling air through the center.

A pressboard sheet $\frac{1}{16}$″ thick separates each pole piece from the yoke to which it is attached to prevent short circuits between laminations.

The shape of the pole faces was derived by

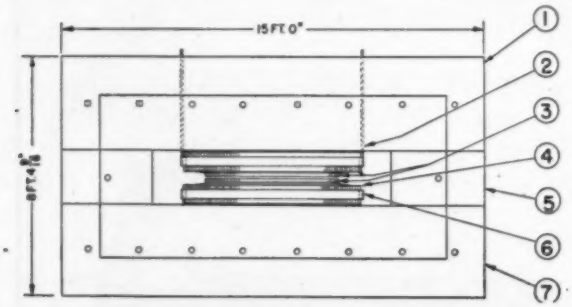


FIG. 4. Magnetic circuit. Wood facing, air ducts, and base omitted.

1. Top yoke.
2. Pole-piece suspension.
3. Centerpiece disks.
4. Pole piece.
5. Leg.
6. Mahogany ring.
7. Bottom yoke.

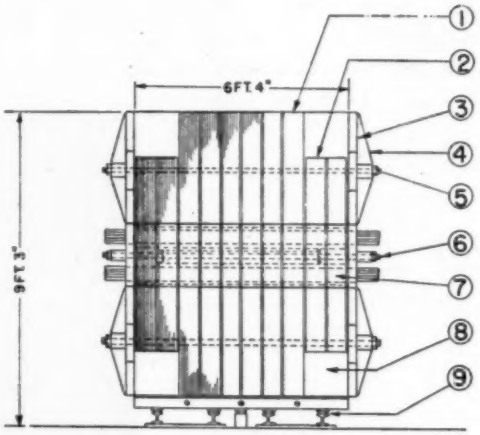


FIG. 5. Side view of magnetic circuit. External air ducts omitted.

1. Large slab.
2. Small slab.
3. Maple board.
4. Anchor.
5. Tie rod through yoke.
6. Tie rod through leg.
7. Leg block.
8. Plywood board.
9. Rubber pad.

means of $\frac{1}{8}$- and $\frac{1}{4}$-scale models made out of solid steel. The measurements on these non-laminated models were all made with direct current excitation which, upon reversal, would produce a ballistic deflection of the galvanometer in the measuring circuit. The shape was chosen so that the magnetic field intensity in the plane midway between the two pole faces varied inversely proportional to the $\frac{3}{4}$ power of the radius. This relationship assures a focusing of the electrons, after the first revolution, to a small beam most of which passes by the electron gun. How well this $\frac{3}{4}$ power law was attained is shown by Fig. 9 in which the logarithm of the field intensity is plotted as a function of the logarithm of the radius.

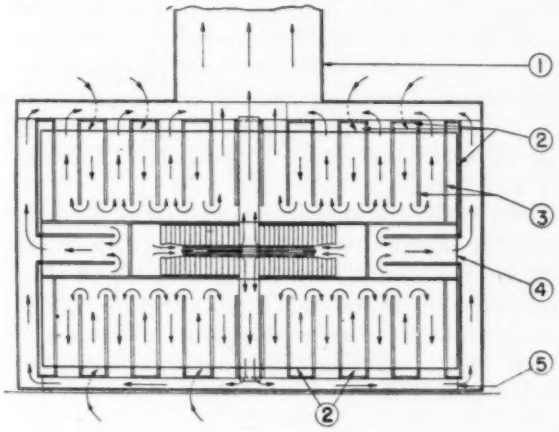


FIG. 6. Cross section of magnet showing forced air cooling.

1. Air passage to penthouse.
2. Air inlet ducts.
3. Wooden spacers.
4. Silicon steel magnetic circuit.
5. Base of machine.

The Centerpieces

The purpose of the centerpieces is to allow fulfillment of the requirement for a circular election orbit, namely, that the flux through the orbit should be equal to twice the magnetic flux density at the orbit multiplied by the area of the orbit.

The magnetic field intensity outside the centerpieces or inside the vacuum tube should be fairly uniform along a line parallel to the axis. This is accomplished by using two centerpiece disks so spaced that the top and bottom air gaps are each equal to half the center gap.

With this arrangement the simple and easily produced pole-face profile of Fig. 7 gave the desired field distribution.

The centerpiece laminations were punched to the proper width corresponding to the thickness

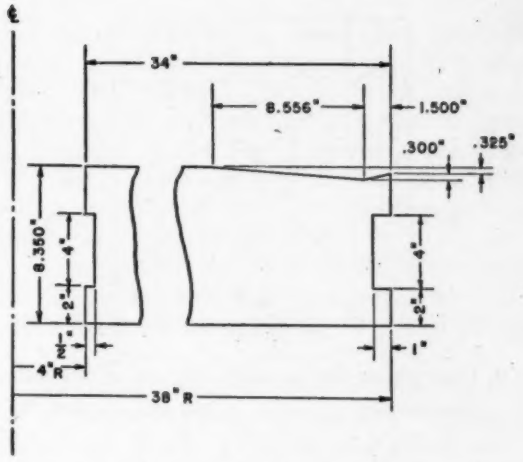


FIG. 7. Profile of pole-piece sector.

of the disks, and the punched strips were cut to various lengths in one-inch steps in order to make up 15-degree sectors.

The centerpieces are supported and spaced apart from one another by rings, and from the pole faces by blocks of Textolite, as shown in Fig. 2. These are placed at such angles to the radius of the pole face or centerpiece that the cooling air is guided at approximately constant velocity through the gaps.

The Yokes

The top and bottom yokes, Figs. 4 and 5, were each made up of six wide slabs flanked by two narrower ones on each side. This arrangement produces a rough approximation to a semi-elliptic cross section accommodating the flux

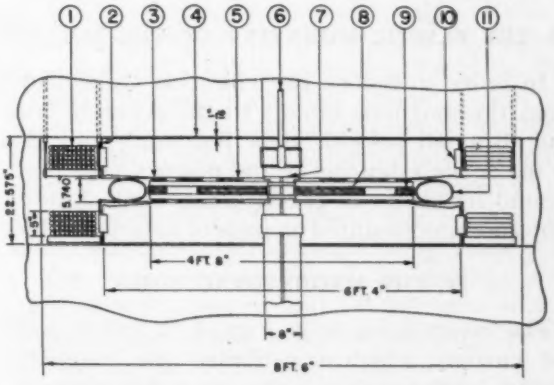


FIG. 8. Enlarged central section of machine.

1. Magnetizing coil.
2. Power input winding.
3. Centerpiece.
4. Pressboard spacer.
5. Textolite spacer.
6. Center suspension.
7. Centering cross.
8. Spacer ring.
9. Retaining ring.
10. Mahogany ring.
11. Vacuum tube.

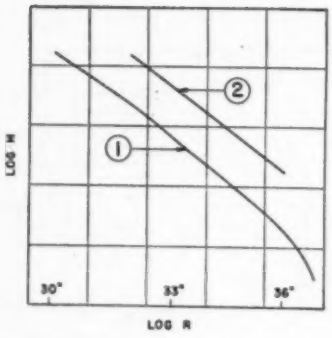


FIG. 9. Logarithmic plot of field intensity *versus* radius.
1. Measured data. 2. Line with $-\frac{3}{4}$ slope.

from the round pole pieces with less variation in flux density and less cross-flux than would result with a simple rectangular yoke cross section.

The top yoke was assembled on a temporary support and the upper pole piece and magnetizing coil were attached to it. This 60-ton assembly was subsequently transferred by means of a crane to its final position.

The Legs

The construction of the legs is simpler than that of the yokes since the steel sheets were of the full size of the blocks. Again six large blocks flanked by two narrower ones on each side were used, thus matching the cross section of the yokes. On top of the legs, along the edge nearest the pole piece, is cemented a strip of pressboard 6″ wide and .045″ thick. The top yoke rests on these two strips and on the Textolite spacers on top of the upper centerpiece.

III. THE ELASTIC MOUNTING OF THE MAGNET

In order to reduce the vibration transmitted from the magnetic circuit to the concrete floor and to avoid possible floor resonance and also to make the vibration of the magnet take place around its center of gravity, the whole 130-ton structure was mounted on pads of natural rubber.

IV. THE MAGNETIZING COILS

The magnetizing coils, Figs. 1, 2, and 8, carry the current which constitutes the magnetomotive force producing the main flux through, and the guiding field at the orbit. Forty thousand-ampere turns r.m.s. are needed for 100-mv operation. This current is distributed over 80 turns of cable, 40 in the top and 40 in the bottom coil. The voltage per turn is approximately 600 at 100 Mv and the reactive voltamperes or apparent power involved therefore amounts to 24,000 kva. This apparent power is furnished at 24,000 volts r.m.s. from a capacitor bank located in an upstairs room in the building. The current is carried from capacitor to coils by cables each at 12,000 volts from ground. For reasons of safety this is a permanent hook-up with no switches or other disconnecting means.

The cable used for the magnetizing coils consists of 61 strands of 0.102″ Formex covered copper wire wrapped with two layers of 0.010″ varnished cambric. The individual wires are insulated from each other to prevent the generation of eddy currents in the copper by the leakage flux, which is perpendicular to the cable. Four layers of varnished cambric provide insulation for the 600 volts between the turns in a single layer. Five turns per layer were used and the maximum voltage between layers is 6000. The adjacent layers are separated by Textolite blocks $\frac{1}{4}$″ thick.

Both coils are supported on radial Textolite bars; the bottom coil is separated $1\frac{1}{4}$″ from the bottom yoke; the bars under the top coil are suspended by brass rods from the top yoke. It is important to provide adequate insulation between input windings, magnetizing coils, and the laminated steel.*

V. THE CAPACITOR BANK

The 24,000-kva capacitor bank which furnishes the apparent power for the magnetizing coils consists of approximately one thousand units,

* Since this article was written, we have had an accidental discharge of the big capacitor to the magnet core, resulting in a complete loss of roentgen ray output. Investigation showed that the magnetic fields at different points around the tube were no longer in phase and that this was caused by an impairment of the insulation between the individual sheets of steel in some of the slabs, caused by the high voltage discharge. The magnetic flux, due to the excess eddy currents in these slabs, was out of phase with that caused by the main magnetizing current. The difficulty has been corrected and the roentgen ray output reestablished by neutralization of the eddy current fluxes by auxiliary fluxes produced by current passing through a single turn of copper strip wound around each of the injured slabs. This experience has served to emphasize the importance in future machines of safeguarding the electrical insulation in the magnetic core. It seems questionable whether the original insulation between sheets can, with adequate respect to the space factor of the steel, be made sufficient to be safe under the high voltage gradients which can be established by an accidental discharge from the capacitor to the core, and that in such machines the core should be effectively shielded from the capacitor circuit, either with solid dielectric or with a grounded conductor.

arranged in 16 angle-iron frames. In each frame the sixty-odd units are connected in parallel. The eight frames on each side of the room are connected in series forming two banks of 24,000 volts and 500 amperes, each permanently connected between the sections of the magnetizing coils as described before. All the frames are mounted on standard 15,000-volt bus bar insulators which rest directly on the floor. The power loss in the capacitor room is $\frac{1}{3}$ of 1 percent of the voltamperes; therefore, 80 kw. This heat is carried off by forced air circulation of 9000 cu. ft. per minute with a temperature rise of the air of 16°C. Baffles are provided above and between the frames to force the air between the capacitors which are spaced $3\frac{1}{2}''$ apart.

VI. LOCATION OF THE ELECTRON ORBIT

The electron orbit was designed to have a radius of 33 inches with a flux density at the orbit of 4000 gauss at 100 Mv.

The actual location of the orbit was determined by two different methods. The first method has been previously described.[7] It employs a set of three coils having diameters comparable to that of the orbit and depends on the fact that the electric gradient is a minimum at the orbit.

The second method is based on the fundamental requirement that the flux through the orbit equals the flux density at the orbit multiplied by twice the area of the orbit. In this method, the voltage generated in a single turn 66 inches in diameter was bucked against the voltage generated in two 1000-turn coils with an effective diameter of $66/\sqrt{1000}$ inches. A cathode-ray oscilloscope was used as the indicating instrument. A null reading should be obtained on the oscillograph if the orbit was located at the 66-inch diameter and if the small coils were placed with their centers on this diameter.

If the orbit however deviates from the desired location, the position of the small coils can be changed until a null reading is obtained; and the actual orbit diameter can then be calculated from the radius at which the small coils produced the null reading. Computation showed that the method allowed determination of the orbit within $\pm\frac{1}{16}$-inch on a radius of 33 inches. The second method checked the three-coil method.

If the orbit at the first assembly of the machine does not have exactly the right diameter a correction is made by lifting the top yoke and changing the spacers on top of the legs and upper centerpiece. Increasing their thickness will weaken the flux through the orbit more than the field at the orbit and will therefore result in a smaller orbit. The required increase in spacing can be calculated for any machine in terms of the error in the orbit radius. For the present machine this calculation showed that for an orbit with a radius 0.100 inch too large the top pole piece should be raised 0.0035 inch.

After the correct orbit location was assured the top yoke assembly was removed and the tube lifted off its assembly ring and placed in its final position, as shown in Fig. 2.

During operation of the machine the process of orbit contraction brings in a further complication. If the primary voltmeter and the megavoltmeter are calibrated in terms of the value of H and R at the normal orbit which produce certain electron energies, then a correction may have to be made for the energy of the electron after orbit contraction, since H and R at the target during contraction are different from H and R at the normal orbit. The correction factor is the product of two other factors: one of which expresses that the field has been strengthened by the orbit contraction (1.0238) and the other of which indicates a change in the product HR owing to change of location only (.9785). The product is a negligible correction factor (1.0023).

VII. VACUUM TUBE AND EXHAUST SYSTEM

The assembled glass vacuum tube (Figs. 2 and 8), is of toroidal shape with an over-all diameter of 74 inches, an inside diameter of 58 inches, a nearly elliptical cross section with an eight-inch horizontal major axis and a $4\frac{7}{8}$-inch vertical axis. It consists of 16 sectors of moulded pyrex glass* having a minimum thickness of $\frac{1}{4}$ inch and heat treated or tempered. Preliminary experiments on single untempered sectors, with a wall thickness somewhat under $\frac{1}{4}$ inch, had shown that on evacuation they would not stand up under the external atmospheric pressure. When made a little thicker and tempered, however, they would withstand 70 pounds per square inch differential pressure, an ample factor of safety.

The technique developed for making these 16 sectors fit together and having them vacuum tight worked out so well that the large number did not prove to be a handicap.

Each glass sector was ground flat on both ends with an angle between the ends of 22.5 degrees. To allow for some distortion taking

*We are indebted to the Corning Glass Works for moulding these curved glass sectors and tempering them.

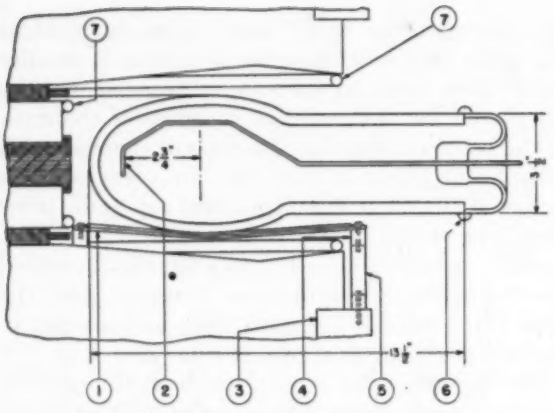


FIG. 10. Target and support of tube.

1. Textolite support strips.
2. Tungsten target, .100" diameter wire.
3. Mahogany ring.
4. Rubber.
5. Textolite post.
6. Cemented joint.
7. Orbit contraction conductors.

place during the tempering operation, each sector was left long enough so that 1/64" could be ground off of each end after tempering.

To prevent the charging up of the inner surface of the tube with consequent disturbance of the electron orbit it was necessary that the inner surface of the sectors should be made conducting. This was accomplished by first sandblasting and then silvering the inside. A chemical silvering process was employed in which solutions of low concentration and definite temperature were used, thus making possible the close control of resistivity.

To make the tube assembly, the sectors were placed on an accurately machined reinforced steel ring with a flat ground top surface. Mounting posts and holding clamps were used to space the glass away from the steel ring and to hold it in place. The flat ends of the glass sectors were placed against each other and feeler shims as thin as 0.5 mil were used to check the line-up. Red Glyptal paint was then applied at each joint to the outside surface only and baked by means of a surrounding heater. The sixteenth sector, required to complete the tube, had to have the angle of one end slightly modified to correct for the cumulative errors in the rest of the assembly.

The ends of the tubulations of the three special sectors were ground square and flat to facilitate the connections to be made. These are butt joints between brass and glass or between glass and glass. For ease of removal, a cement made of pinetar and shellac was used to cover these joints. Fig. 10 shows detail construction of the target.

The connection to the pump system is rigid and consists of a $3\frac{1}{2}$ inch glass elbow passing from the tube to the liquid air trap. This trap is of the concentric type, the inner member being $3\frac{1}{2}$ inches and the outer tube $5\frac{1}{2}$ inches in diameter. Between the liquid air trap and the vacuum tube are attached standard glass bulb ionization gauges. The liquid air trap is connected to a mercury vapor pump which has a speed of 80 liters per second. Ahead of the vapor diffusion pump is located a vacuum valve sealed against outside leakage by the use of a metal bellows. This valve is kept closed whenever the tube is not soon to be put in operation. It connects to a rotary rough vacuum pump, having an automatic centrifugal shut-off valve which keeps oil or air from leaking into the vacuum system whenever the motor stops.

Liquid air is at all times kept around the trap. When the electron gun or the target is to be replaced for experimental work, carefully dried air is let into the tube at a seal-off tip located between trap and tube. After the tube system is again closed and the pumps turned on, the pressure will drop to 0.01 micron within 1 hour and the tube can then be operated satisfactorily. By the end of another hour the pressure will usually come down to 0.001 micron or less and can without difficulty be maintained at this level during continuous operation of the accelerator.

Because of the very rigid connection between the tube and the pump system, considerable apprehension was at first felt lest too much vibration should be communicated to the glass from the magnet. This was prevented however by a rather elaborate system of elastic mounting for the tube. This consists of 19 flexible supports, each of which (Fig. 10) is made up of two Textolite strips, $\frac{1}{8}'' \times 2\frac{1}{2}'' \times 11''$ laid horizontally one on top of the other with the ends resting on $\frac{1}{2}''$ soft rubber cubes, one cube near the centerpieces, two cubes mounted on a Textolite block screwed to the mahogany ring around the pole piece. This support bends approximately $\frac{1}{4}''$ under the weight of the tube and has a natural frequency of approximately five cycles per second. The rubber cubes are pierced and are mounted on pins which allow flexibility without shift. The rubber serves, therefore, to furnish flexibility in horizontal directions. The elastic tube support is so effective that a hand placed upon the tube feels no vibration when the exciting current is turned on the magnet.

VIII. ELECTRICAL CIRCUITS

Figure 11 shows the main power circuits. A motor generator set of large enough three-phase rating was available so that the 200-kw single phase load could be handled without danger of pole-tip overheating because of eddy currents. The advantage of the use of a motor generator set over direct power line connection is the independence of line voltage fluctuations and accuracy of voltage adjustment by means of a voltage regulator acting upon the field of the exciter of the generator field. The motor is synchronous and the accurately maintained power line frequency is therefore also present on the output side of the generator. However it was found that the small frequency variations of ± 0.1 cycle per second that do occur in the power frequency caused corresponding but relatively large variations in accelerator voltage if the generator was connected directly to the input winding (5 and 6, Fig. 11) of the accelerator. The explanation of this effect lies in the considerable variation of the power factor of the load with frequency and the corresponding voltage drop caused by the synchronous reactance of the generator. When the frequency is 60 cycles per second and the correct number of capacitors is connected in the capacitor bank, the input power factor is unity. However, when the frequency increases by 0.1 cycle per second the capacitor-bank kva increases by $0.1/60 \times 24{,}000$ or 40 kva and the magnetizing kva decreases by 40 kva resulting in a load on the generator with an 80-kva leading component corresponding to 33 amperes at 2400 volts. With a generator synchronous reactance of nine ohms this additional current would tend to boost the output voltage by a vector component of 300 volts and if the frequency variation occurs rapidly, the voltage regulator will not be able to correct the voltage. The obvious solution is to place a series capacitor in the circuit which just neutralizes the synchronous reactance. It was found that this solution worked perfectly in preliminary work in which the accelerator was operated at 50 Mv by connecting the four turns of the input power winding around the top pole piece in series with the four turns of the bottom pole piece and operating the generator at full voltage. In this case the reactive kilovolt amperes involved were only one quarter of those at 100-Mv operation. When the input power windings were connected in parallel for 100-Mv operation the system involving the series capacitor became

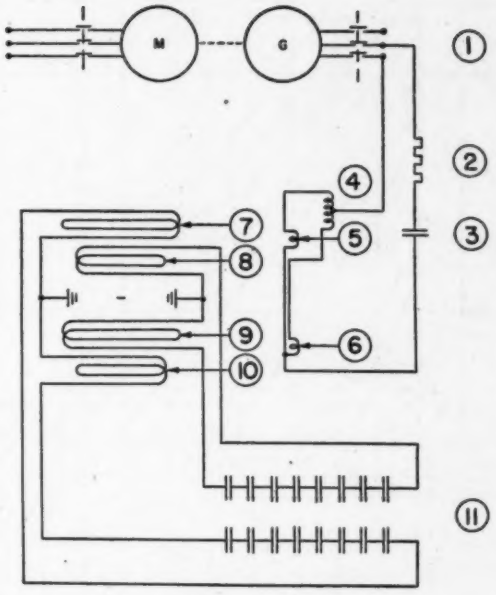


FIG. 11. Main power circuits.

1. Motor generator set, 1200 kva, 2400 volts.
2. Stabilizing resistor.
3. Series capacitor.
4. Equalizing transformer.
5. Four-turn input power winding around top pole piece.
6. Four-turn input power winding around bottom pole piece.
7. Twenty-turn outside section of upper magnetizing coil.
8. Corresponding inside section.
9. Outside section of lower magnetizing coil.
10. Corresponding inside section.
11. Capacitor bank rated 24000 kva. Each symbol represents a rack with capacitors rated 3000 volts.

unstable, power oscillations of low frequency taking place from the motor end of the system to the tuned circuit with its 24,000 kva. It was found that these oscillations could be completely suppressed by an eight-ohm resistor dissipating between 50 and 80 kw depending on the tuning. The tuning varies from day to day and during operation on account of the temperature coefficient of the capacitors. Even though this coefficient is only -0.0005 per degree C, a change of 1 percent can be expected with a temperature range of 20°C from a cold start on a cold day to the final temperature on a hot day. This change corresponds to a 240-kva total decrease in capacitor kva. The circuit is tuned once for all for the medium operating temperature so as to give unity power factor for the power input. The equipment can then be operated over the full range without retuning.

The injector and orbit contraction circuits are similar to those used for the 20-million volt accelerator.[10,11]

[10] D. W. Kerst, "A 20-Million Electron-Volt Betatron or Induction Accelerator," Rev. Sci. Inst. **13**, 387–394 (1942).
[11] D. W. Kerst, "The Betatron," Radiology **40**, 115–119 (1943). H. W. Koch, D. W. Kerst, and P. Morrison,

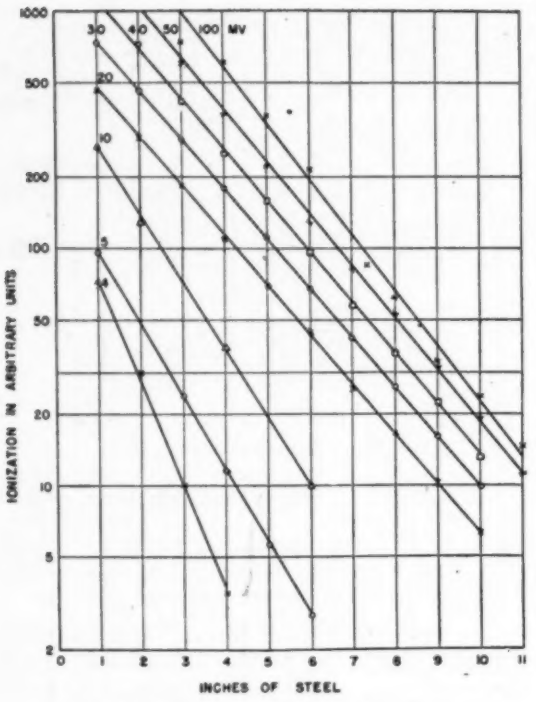


FIG. 12. X-ray transmission through steel.

Ordinates in arbitrary units of intensity without relation between curves. 100-, 50-, 40-, 30-, 20-, and 10-megavolt x-ray curves taken with direct reading ionization chamber. Five- and four-megavolt curves taken with fluorescent screen and photo-multiplier tube.

IX. THE COOLING SYSTEM

In a penthouse above the machine room is located a ten-hp blower capable of delivering 8000 cubic feet of air per minute at one inch of water pressure drop for dissipating the iron loss of 100 kw in the magnet. The air is passed through filters before it enters the room. Three sets of louvers are thermostatically controlled and determine the ratio of recirculation to outdoor intake and exhaust. The low pressure end of this penthouse is connected through a passage to the top of the machine. There it receives the air from the external duct system built around the silicon steel magnetic circuit. Part of this "external" duct system is also formed by the base of the machine. Figure 6 shows a cross section through the machine through one of the $\frac{3}{8}$" wide cooling slots between the slabs of the yokes and the blocks of the legs. The rectangular air inlet ducts extend from the front to the back of the machine and can be seen in the photograph Fig. 1 where they show up as dark openings. In Fig. 6 the air guiding maze formed by the soft wooden spacers between the silicon steel slabs can readily be followed. The bottom of the top yoke and similar sides of legs and bottom yoke are completely closed by wooden strips so that air is forced throughout as indicated by the arrows. The eight-inch holes in the center of yokes and pole pieces allow passage of air from the space external to the vacuum tube through the top and bottom gaps to the main duct.

The magnetizing coils receive no forced air cooling, the $\frac{1}{4}$" vertical slots between the layers providing ample cooling by convection.

Three places on the vacuum tube receive additional cooling from the low pressure shop air line. These are the cemented connections of the vacuum system, the electron gun, and the target.

The capacitor room has its own air circulating system located in the penthouse above, with temperature controlled louvers and outdoor intake and exhaust. The five-hp blower of this system is capable of circulating 9000 cubic feet per minute at $\frac{1}{2}$ inch of water pressure drop. With 80-kw loss in the capacitors corresponding to $\frac{1}{3}$ percent of 24,000 kva at 100-mv operation, the temperature rise of the air in steady-state conditions will be 16°C.

X. THE BUILDING

The required x-ray protection, the concentrated floor load, and the special crane facilities needed for the accelerator made it necessary to

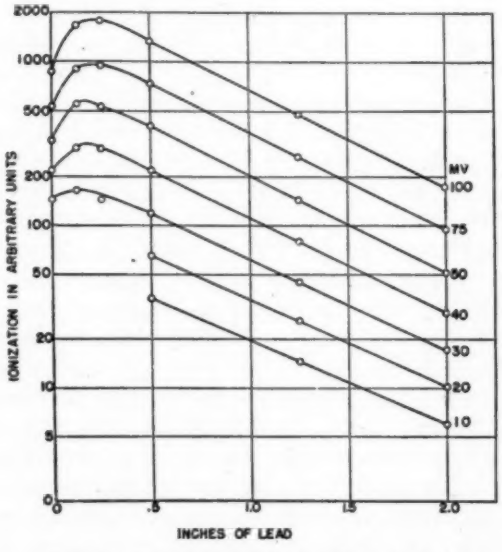


FIG. 13. X-ray transmission through lead.

Ordinates in arbitrary units of intensity without relation between curves. 100-, 75-, 50-, 40-, 30-, 20- and 10-megavolt x-ray curves taken with direct reading ionization chamber.

"Experimental Depth Dose for 5, 10, 15 and 20-Million Volt X-rays." Radiology 40, 120–127 (1943).

construct a special building for the machine. The floor plan of the building, Fig. 3, shows the wall thickness of the accelerator room to be at least 36 inches. This thickness of the concrete walls is maintained up to a height of 15 feet where the walls have a ledge and step back to brick and glass blocks. Since the x-ray beam is relatively narrow and scattered radiation from the tube is blocked by the top yoke this height was found to be sufficient. The ledges of the concrete walls were used for the rails of the 60-ton crane. The floor of the machine room is reinforced concrete 16 inches thick. The massive walls of the building and the nature of the soil made it necessary to drive piles to prevent settling. A motor-operated one-inch thick steel door is provided between the hallway and the machine room principally to stop primary and scattered electrons.

As indicated in the floor plan there is, in that portion of the wall exposed to the x-ray beam, a section built up of loose concrete blocks. In case it is desired to utilize the beam outside the building a portion of this section can be removed.

In the control room are located the panels containing the operating handles, switches and meters of the main power circuit, the injector circuit, the orbit contraction circuit, and the ionization vacuum gauge. Other instruments, such as the indicating instruments of direct reading ionization chambers and a Geiger-Mueller counter, are also located in sight of the operator. A megavolt meter gives a direct reading of the electron energy at the instant of orbit contraction.

The blowers for the machine and capacitor rooms are started from the control room. An airflow-operated vane switch provides an interlock which allows the main power to come on only after the capacitor room blower is started. The machine room blower can be started either manually by push button or automatically by means of a relay which picks up at half-generator output voltage. Interlock switches are placed on all doors to the machine and capacitor rooms. In addition, the door to the capacitor room is kept locked for safety of personnel.

The capacitor room is located above the control room and extends over the hallway. Openings are left in the ceiling at each end, connecting the capacitor room to the corresponding blower and filter system in the penthouse. Similar slots are provided in the machine room ceiling connecting to the machine air cooling system in the penthouse.

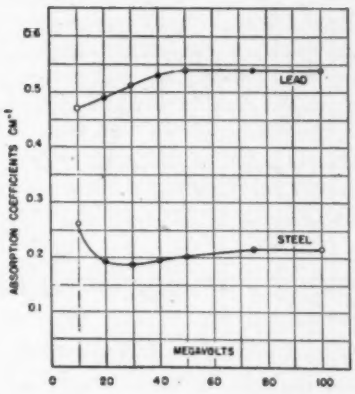


FIG. 14. Linear absorption coefficients of x-rays in lead and steel.

The coefficients in reciprocal centimeters were computed from the slopes of the curves of Figs. 12 and 13.

The noise level in the accelerator room during operation is very high, between 110 and 120 decibels, partially because of the nature of the entirely closed room with concrete walls and glass brick windows.

XI. SOME EXPERIMENTAL RESULTS

The absorption curves for steel and lead for various x-ray energies are shown in Figs. 12 and 13. The energy of the x-rays was adjusted by means of the phase shifter of the orbit contraction circuit. The slopes of the curves of Figs. 12 and 13 are shown in terms of the linear absorption coefficients in Fig. 14.

The x-ray distribution for different energy levels is shown in Fig. 15. The cone of x-radiation narrows very rapidly in going from 20 to 50 Mv. There is less change in the cone between 50 and 100 Mv.

These data were taken with an ionization chamber with an effective diameter of 0.85 degrees limiting the resolution. For this reason a more sensitive radiographic method was devised to determine the cone of radiation at 100 Mv.

The radiographic determination of the half-value beam diameter at 100-Mv operation was based on the simultaneous exposure and development of two films placed respectively at 8 ft. and $8\sqrt{2}$ ft. (11′ 4″) from the target in the beam. The half-value beam diameter at 8 ft. was taken to be the diameter on the film at which the same density was found as at the point of maximum

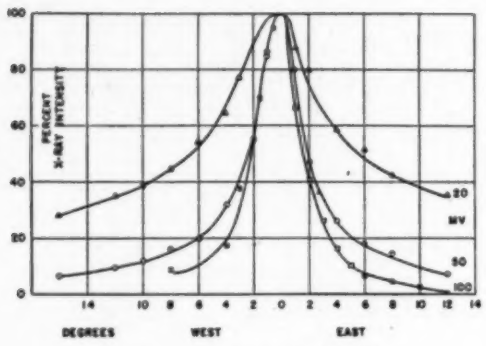


FIG. 15. X-ray beam distribution.

The data were taken with the center of the chamber at 19 feet 1 inch from the target and are plotted for 20-, 50-, and 100-megavolts in percent of the maximum of each beam. The chamber had the effective diameter of $3\frac{3}{5}$ inches or 0.85 degrees, limiting the resolution. Beam diameters at 50 percent as read from the curve are, respectively, 12.3, 4.2, and 3.7 degrees.

density on the film exposed at $8\sqrt{2}$ ft. distance. This value corresponded to 2.0 degrees.

The x-ray intensity in the center of the beam is the greatest at 100-Mv operation. The output drops as the phase of the orbit contraction pulse is advanced and the energy of the electrons impinging on the target is decreased. Figure 16 shows how the x-ray intensity varies as a function of the x-ray energy. In one of the curves, $\frac{1}{4}$-inch lead was used as an added filter in front of the ionization chamber and in the other curve, this lead was omitted and .005-inch aluminum was the only filter. With this thickness of added lead filter the intensification factor offsets the absorption factor at 30 Mv and above.

Figure 17 is a radiograph taken with 100 Mv x-rays. In this radiograph made at a 20-foot target-film distance, a clock was placed on the source side of 4 inches of steel. The dark spot appearing on this film is due to the concentration of the x-radiation coming from the machine in the 2.0 degree half-value cone at 100 Mv.

The x-ray output at 100-Mv operation, as measured with a Victoreen 100R thimble chamber placed within a $\frac{1}{4}$-inch lead jacket at a distance of 550 cm from the target and computed to the standard 100-cm target thimble distance, amounted to 2600 Roentgens per minute.

XII. PULSE OPERATION FOR CLOUD CHAMBER WORK

For cloud-chamber studies the machine need not be run continuously since only a single burst of x-rays of short time duration is desired at approximately one minute intervals.

For this use the machine was reconnected as shown in the diagram of Fig. 18. The capacitor bank is split in the center; one-half can be charged to a positive potential by means of one kenotron and the other half to a negative potential by means of another kenotron and the same high voltage transformer. An oil circuit breaker serves to complete the circuit and to start an oscillatory discharge with a time constant of 0.28 seconds. The breaker never has to open the circuit under current and a standard 13.8-kv oil circuit breaker could be used even though the d.c. potential difference may be as much as 40 kv before closing.

A 250-lb. weight keeps the breaker normally closed, which is the "safe" condition for the circuit to be permanently maintained. A compressed air operated piston serves to push up the weight and open the breaker. When deenergized a solenoid operated air valve allows the air to escape and the breaker to close. Relays shut off the input power of the charging circuit when the breaker closes to prevent unnecessary loading of the kenotrons.

Since all other connections remain intact, no heavy leads or cables need be reconnected and the machine can be switched instantaneously from continuous operation to pulse operation or *vice versa* by simply throwing an interlock-transfer switch. During the continuous operation the oil circuit breaker simply passes the 1000-ampere 60-cycle current of the capacitor bank.

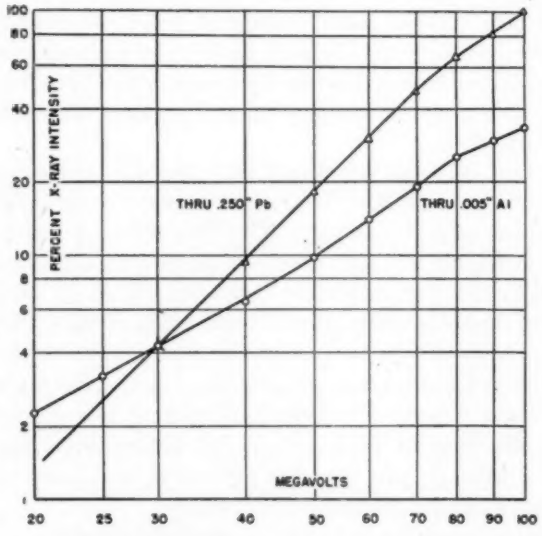


FIG. 16. Logarithmic plot of x-ray intensity *versus* megavolts.

Data obtained with direct reading ionization chamber in center of beam at 150 cm from target. Effective diameter of chamber 9.5 cm or 3.6 degrees. The point of intersection at 30 megavolts of the two curves was obtained by having a .250-inch thick sheet of lead swing in pendulum form in front of the chamber and adjusting the megavolts until the ionization meter reading remained constant.

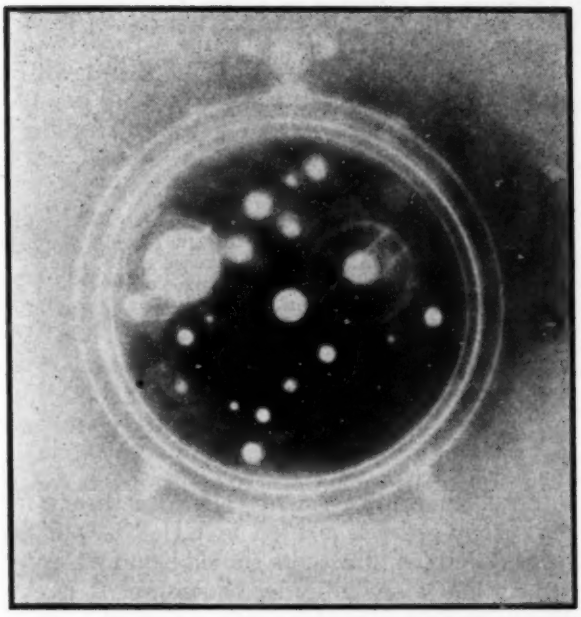


FIG. 17. Radiograph through 4 inches of steel taken with 100-million volt x-rays at 20-foot target-film distance.

A clock was placed on the source side of the steel block. The dark spot on this radiograph is owing to the concentration of the x-radiation coming from the electron accelerator in the 2.0 degree half-value cone at 100 Mv.

For pulse operation for cloud-chamber work various combinations are possible, such as:

1. Push button or motor driven cam control of the pulse circuit and coincidence counter-control of the cloud chamber. This has the advantage that the cloud chamber is electrically independent from the accelerator and can be similar to the equipment used for the study of cosmic rays.

2. The cloud-chamber cycling mechanism initiates the trip of the pulse circuit and various delay circuits trip the chamber and flash the light for photographic recordings. In this case it is important to use some electrical pulse from the accelerator such as the orbit shift pulse to flash the light through a delay circuit in order to get accurate timing.

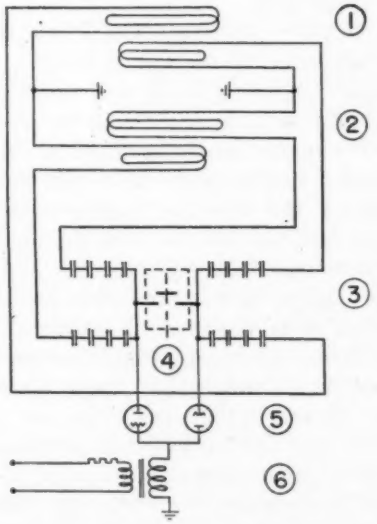


FIG. 18. Circuit for pulse operation.

1. Magnetizing winding around upper pole piece of accelerator.
2. Magnetizing winding around lower pole piece of accelerator.
3. Capacitor bank.
4. Oil circuit breaker, gravity closing, air pressure opening.
5. Kenotrons.
6. High voltage power transformer.

It was found that the electron injection and orbit shift circuits were instantaneous in their action and worked as well with pulsing as continuously. This is caused in part by the fact that the polarity of the charge on the capacitors was made so as to make the second and not the first half-cycle the period of electron acceleration thereby allowing charging of the capacitors in the injector and orbit shift circuits during the first half-cycle. Without further precautions, a train of about 12 x-ray pulses is generated which may or may not interfere with cloud-chamber work, in which only the first one is desired. For this reason a relay is provided which breaks the injector charging circuit after the first half-cycle, thus permitting only one electron injection and one pulse of x-rays.

The pulse type of operation is very quiet compared to continuous operation since during less than one second out of each minute the machine can be heard. The intensity of x-radiation required for cloud-chamber work is extremely low and the occurrence of only one pulse per minute allows the cloud-chamber operator to be present near his equipment when necessary for visual observation and adjustments.

Because of its lack of noise during pulse operation and the sudden application of high voltage, direct access to the electron accelerator must be blocked by electrically interlocked safety fences if and when personnel are permitted to be in the same room with the equipment during pulse operating procedure.

ACKNOWLEDGMENT

The authors wish to make grateful acknowledgment to Dr. W. D. Coolidge for his continued interest and many helpful suggestions given through the course of this work. We are also indebted to many of our associates in the Research Laboratory for their valuable assistance in the course of construction of this new research tool.

Thermal Expansion and Second-Order Transition Effects in High Polymers

PART II. Theory

R. F. Boyer and R. S. Spencer
Physical Research Laboratory, The Dow Chemical Company, Midland, Michigan

(Received May 12, 1945)

The nature of the thermal expansion anomaly in high polymers, known as the second-order transition, is examined in some detail. It is suggested that below the transition temperature, *Tm*, polymer chains can expand sidewards but not parallel to their length. At *Tm* lengthwise expansion becomes prominent, thus accounting for the sudden increase in thermal expansion. Experimental results are presented showing the anisotropic expansion of oriented polymers below *Tm*. The transition effect is then treated as a problem in viscous flow, which gives rise to various semi-empirical plots connecting *Tm* with applied forces, plasticizer content, and time effects. The brittle point, T_b, involves highly elastic deformation, and is shown to be a fundamentally different test, although *Tm* and T_b are sometimes numerically equal. Various factors influencing the brittle point are reviewed briefly.

INTRODUCTION

THE term second-order transition, as used here in connection with high polymers, refers to the discontinuity in thermal expansion, specific heat and other physical properties which occurs over a relatively small temperature range characteristic of any given plastic composition. The first paper[1] in this series presented some typical data of the same kind that has been observed by numerous investigators. We now wish to present some theoretical aspects of this problem, primarily by discussing both a mechanism of thermal expansion for long chain compounds, and the effects of molecular weight, external forces, plasticizers, and heating rates on the value of the transition temperature, *Tm*, and on the magnitude of the difference $\Delta\beta$, in cubical expansion coefficient above and below *Tm*.

An approach to this problem in terms of first principles is apparently faced with tremendous difficulties. Kirkwood has treated second-order transitions for the hydrogen bromide case,[2] and also for fatty acid monolayers which exhibit second-order transitions in the pressure *versus* area relationships.[3] His work predicts the qualitative nature of the effect, but the approximations used in the interest of simplicity rule out any rigorous quantitative conclusions. Onsager[4] has more recently treated second-order transitions for some special cases, such as that of a two-dimensional crystal. Here again the author calls attention to the mathematical complexity of the problem. Slater[5] has developed a theory for the second-order transition in KH_2PO_4, but his theory predicts that there should be a first-order rather than a second-order transition. It is evident that there are still further obstacles introduced by the long chain nature of a polymer molecule. London[6] has discussed the general aspect of the force problem for highly elongated molecules. There are, however, many things which can be discussed qualitatively in terms of a specific mechanism for the thermal expansion of an articulated, anisotropic molecule, and quantitatively in terms of the melt viscosity behavior of polymers. The closely related problem of brittle points follows as a logical development of these ideas.

There is also a thermodynamical approach to the second-order transition question, whereby mechanism can be ignored and certain general criteria established for describing the phenomena. Ehrenfest[7] has furnished the classical solution for low molecular weight materials, while Ueberreiter[8] has outlined the various energy-entropy considerations appropriate to polymers. More recently Pines[9] has offered thermodynamic arguments to explain how a second-order transition might occur over a range of temperature. One apparent difficulty in applying thermodynamics to the polymer case is that these second-order transition effects in high polymers are predominantly rate and not equilibrium phenomena. It has, in fact, been suggested that if a thermal expansion experiment were carried out slowly

[1] R. F. Boyer and R. S. Spencer, J. App. Phys. **15**, 398 (1944).
[2] J. G. Kirkwood, J. Chem. Phys. **8**, 205 (1940).
[3] J. G. Kirkwood, *Surface Chemistry* (Pub. No. 21, AAAS, Washington, D. C., 1943), p. 157.
[4] L. Onsager, Phys. Rev. **65**, 117 (1944).
[5] J. C. Slater, J. Chem. Phys. **9**, 16 (1941).
[6] F. London, J. Phys. Chem. **46**, 305 (1942).
[7] P. Ehrenfest, Comm. Leiden, Suppl. 75B (1933).
[8] K. Ueberreiter, Zeits. f. physik. Chemie **B45**, 361 (1940).
[9] B. J. Pines, J. Exp. Theo. Phys. U.S.S.R. **9**, 963 (1939).

enough, there might be no break in the volume temperature curve.

Alfrey, Goldfinger, and Mark[10] have emphasized the time aspect of second-order transitions in polymers by stating that they merely represent the temperature at which the relaxation time for some internal mechanism is equal to the time scale of the experimental method of observation. These authors then logically conclude that such effects are not at all concerned with second-order transitions in the strict thermodynamic sense of that term, and should not be called such. Previous to their work, Jenckel[11] had made time studies of volume change in selenium glass and colophony resin. He was able to lower the second-order transition temperature of the former 7°C by slowing down the rate of heating. Richards[12] has pointed out that *Tm* for glucose glasses increases from 7° to 77°C as the rate of observation is speeded up in the sequence heat capacity-thermal expansion—1000 cycle dielectric constant. A treatment of second-order transitions in terms of the viscous flow of polymer chains provides a logical interpretation of these time effects.

In addition to the stress on time effects, the other viewpoint which has emerged consistently concerns the role of free rotation about carbon-carbon bonds. The second-order transition temperature is thought to be that temperature at which free rotation begins; or perhaps more accurately, where it occurs frequently enough to accommodate the particular process being observed. This idea is particularly evident in the writings of Fuoss[13] on the electrical properties of polymers. Alfrey and Mark[14] have also expressed this view, citing in part the work of Muller[15] on second-order transitions in paraffins, where the entire molecule begins to rotate about its long axis on approaching the transition temperature. Tuckett[16] has written very explicitly concerning segment rotation and second-order transitions, while Jenckel[17] has offered additional evidence favoring this concept. A viscous flow mechanism for transitions in polymers, such as will be attempted in the present paper, is in accord with, and in fact requires, the concept of free rotation. With these introductory remarks we can now examine the nature of the thermal expansion process in polymers.

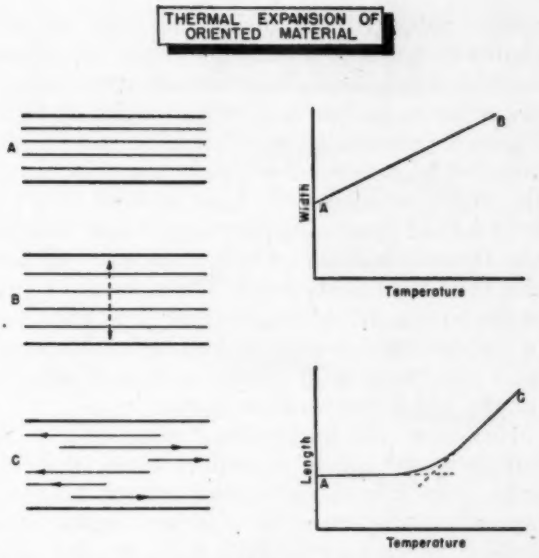


FIG. 1. Idealized representation of the anisotropic linear thermal expansion of an oriented polymer.

MECHANISM OF THE THERMAL EXPANSION PROCESS IN HIGH POLYMERS

It has been found on several occasions[10,11] that when the temperature of a high polymer is changed a certain time is necessary for the volume of system to reach equilibrium at the new temperature. Very little has been said, however, as to the internal mechanism giving rise to this lag behind equilibrium which occurs below the transition point. Alfrey, Goldfinger, and Mark[10] speak of segment diffusion and the diffusion of holes as giving rise to the time effects which have been noted, but no one has hitherto used these concepts to explain and correlate the diverse phenomena reported in the literature. It is our hope that the simplified mechanism we are about to propose may serve to provide a more unified interpretation and clearer understanding of the various transition phenomena observed.

When a simple organic material expands with increasing temperature its constituent molecules move farther apart in all directions as a consequence of their enhanced thermal energy. In the case of a long chain compound adjacent atoms along the chain are bound together by primary valence forces which constrain the chain into acting as a unit or as a sequence of segmental

[10] T. Alfrey, G. Goldfinger, and H. Mark, J. App. Phys. **14**, 700 (1943).
[11] E. Jenckel, Zeits. f. Electrochemie **43**, 796 (1937).
[12] W. T. Richards, J. Chem. Phys. **4**, 449 (1936).
[13] R. Fuoss, J. Am. Chem. Soc. **63**, 369 (1941).
[14] T. Alfrey and H. Mark, Rubber Chem. and Tech. **14**, 525 (1941).
[15] A. Muller, Proc. Roy. Soc. **A127**, 417 (1930); **A158**, 403 (1937); **A178**, 227 (1941).
[16] R. F. Tuckett, Trans. Faraday Soc. **38**, 310 (1942).
[17] E. Jenckel, Kolloid. Zeits. **100**, 163 (1942).

units. Sideward expansion between adjacent chains should obey normal laws since the cohesive forces holding chains together are approximately the same as in low molecular weight materials. Figure 1 presents an oversimplified picture of an oriented high molecular weight material. In (*A*) the chains are arbitrarily lined up with their ends in two fixed planes. Application of heat increases the thermal motion of the molecular segments and they tend to separate. The sideward motion of the chains should be relatively free, the chains on the average moving farther apart, as in (*B*). This should result in a more or less straight line for the width-temperature curve.

However, the longitudinal expansion of the sample would appear to require a special mechanism. Any increase of carbon-carbon bond distances, or distortion of valence angles seems unlikely in view of Muller's x-ray studies of the expansion of simple paraffins.[15] He showed that from liquid air temperature to the melting point of the paraffins there was no change in identity period along the chain, at least within the limits of error of his measuring technique. A question arises concerning the possible role of chain uncoiling in providing for thermal expansion parallel to the chains. The same phenomenon of free rotation necessary for viscous flow is also the basis for rubber-like elasticity connected with the coiling and uncoiling of polymer chains. It may well be that the first step in the lengthwise motion of a polymer chain is one of uncoiling. However, Poisson's ratio for high polymers is very close to 0.5, at least when they are in a rubber-like state, i.e., highly plasticized or above the transition temperature.[18,19] It is not known precisely how close to 0.5 this ratio would be for a material like polystyrene near its transition temperature, but it would seem that uncoiling or stretching of chains would have little effect on the volume. In one experiment, a sample of polystyrene elongated at an elevated temperature to five times its original length and then quenched had increased in density from 1.0408 to 1.0468. Certainly in the experiments we are about to consider the chains are already highly uncoiled and the net tendency of any temperature increase is to make them want to coil up.

The lengthwise expansion thus seems to require a gross displacement of entire chains or chain segments parallel to their length. This in turn seems to require viscous flow. Two factors are in opposition: The thermal energy of the chain segments tends to cause length-wise flow, while the internal viscosity (sideward attraction between chains) retards the flow. These two factors result in a velocity which characterizes the expansion along the orientation axis at a given temperature. As the temperature increases the viscosity decreases, i.e., the chains move farther apart in the plane normal to the orientation, while at the same time the thermal energy increases. Hence the velocity of flow becomes greater. For temperatures well below the transition temperature this velocity is so small that no change should be observable in the length of the oriented specimen during the time interval involved. Let us assume that the chains are initially at equilibrium, i.e., there is no tendency for them to flow at constant temperature, and let us then increase the temperature at a uniform rate. As the temperature rises, the flow velocity of the segments increases, and they are displaced from their equilibrium positions at an ever increasing rate. The rate of heating for this part of the curve is such that the velocity of the chain flow is not great enough to enable the chains to arrive at equilibrium at any temperature after heating begins. At first, this displacement of the chains is too small to be observed and the sidewise expansion alone shows up in the volume-temperature curve. Hence, we get a straight line. Eventually, however, the ratio of the displacement from the initial equilibrium position to the chain

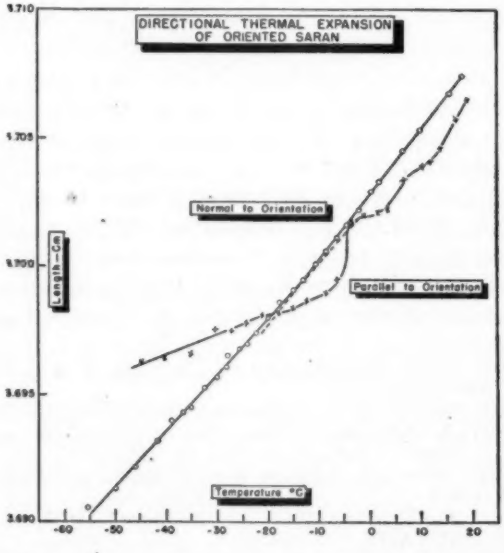


FIG. 2. Linear thermal expansion of oriented saran parallel and at right angles to the direction of orientation.

[18] W. W. Voght and R. D. Evans, Ind. Eng. Chem. **15**, 1015 (1923).
[19] W. L. Holt and A. T. McPherson, J. Research Nat. Bur. Stand. **17**, 657 (1936).

length becomes equal to the ratio of the smallest observable volume increment to the total volume, and the curve begins to deviate from a straight line. This is just another way of saying that the displacement of the chains finally produces a volume change which is greater than the observational error. The deviation increases with rising temperature until we reach a point where the chains have finally come to their equilibrium positions corresponding to the temperature at that point. At every temperature above this the chains have had more than enough time to reach new equilibrium positions while the system has been heating and approaching the given temperature. The expansion is now like that of a liquid and we again have a straight line for the volume-temperature curve, at least over a small temperature interval. The expansion coefficient has increased, however, over that of the lower straight-line portion, and there is an apparent transition as indicated in Fig. 1*C*.

This proposed mechanism of thermal expansion in polymers immediately suggests several predictions. The first one concerns an expected anisotropy in the expansion of oriented materials. A second prediction is that for unoriented specimens the normal sideward expansion of the chains below Tm will be hindered because of their inability to slip lengthwise. Hence the over-all volume expansion below the transition temperature should be less than the volume expansion of an oriented sample, but equal to it above the transition point. In other words, $\Delta\beta$ would be greater in the absence of orientation. The experimental investigation of these predictions will now be discussed.

With most oriented polymeric materials it is difficult to make linear expansion measurements above the transition temperature because contraction occurs from the coiling up of the extended chains. Oriented, partially crystalline saran can be used, however. In this case, the crystalline regions are sufficiently stable for some 50°C above the transition temperature to maintain the specimen in an oriented condition. The constraints imposed on expansion by occasional virtual cross-linking in crystalline regions complicate the drawing of conclusions, but do permit the measurements to be made.

A supercooled copolymer of vinylidene chloride and 14 percent vinyl cyanide was oriented by stretching, allowed to crystallize, and cut into two specimens, one parallel and one perpendicular to the orientation axis. The thermal expansion of

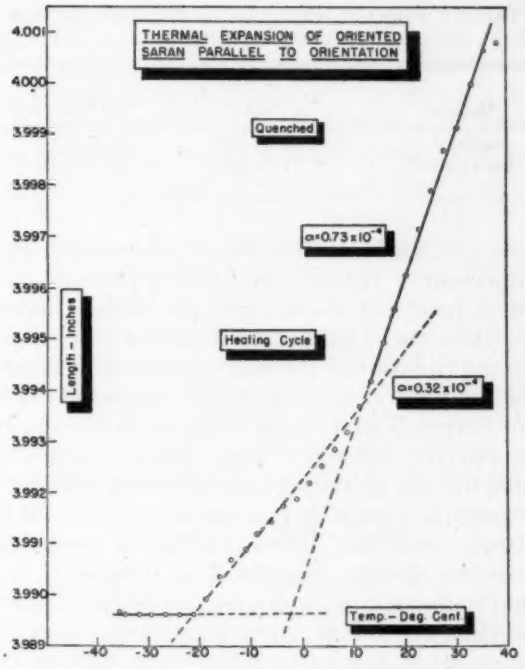


FIG. 3. Linear thermal expansion of a rapidly quenched saran sample parallel to the direction of orientation. Linear expansion coefficients are indicated at several points along the curve.

these samples was then observed in a linear quartz dilatometer (ASTM D696-42T), with results as shown in Fig. 2. It may be seen that the sideward expansion still exhibits a very weak transition at −12°C, but that quite a pronounced transition is present in the lengthwise expansion. The oscillations in the lengthwise expansion curve are believed to arise from stresses frozen in the sample during cooling. A case has been recorded by Morey[20] for glass in which this effect actually amounted to a net contraction of the sample over a temperature interval, followed by an expansion. It was found that for extremely quick chilling of the saran sample the subsequent lengthwise expansion was entirely negligible below a certain temperature, as seen in Fig. 3.

In principle, then, the first prediction was fulfilled, except for what might be ascribed to imperfect orientation of the original specimen or to the secondary influence of crystallinity. A further check was then made on oriented polystyrene, which is non-crystalline. Polystyrene is too soft above the transition temperature to permit reliable measurement of its linear expansion. Moreover, since it loses orientation

[20] G. W. Morey, *Properties of Glass* (A.C.S. Monograph No. 77, 1938), p. 266.

TABLE I. Effect of orientation on the linear thermal expansion of polystyrene.

Material	α Parallel	α Normal (By difference)
oriented	0.54×10^{-4}	0.83×10^{-4}
unoriented	0.775×10^{-4}	0.79×10^{-4}

above *Tm*, such measurements were restricted to temperatures below *Tm*. The technique employed here was to measure the linear thermal expansion coefficient parallel to the orientation axis and the volume thermal expansion coefficient over the same range, the linear expansion coefficient perpendicular to the orientation then being calculated by difference. The experimental justification for this procedure may be found in Table I. The sample designated as unoriented should be isotropic, and the linear expansion coefficient across the sample, calculated by difference, was found to check quite well with the observed linear expansion coefficient in the lengthwise direction. The oriented specimen, however, exhibited the same sort of anisotropy as did oriented saran, although to a smaller degree.

It is interesting to note that Vieweg and Schneider,[21] studying the linear expansion of laminated synthetics, found higher expansivity perpendicular to the laminae than in their plane. For example, a paper-base laminated material had a linear expansion coefficient of 25×10^{-6} parallel to the laminae and one of 140×10^{-6} perpendicular to the laminae. This is a macroscopic phenomenon depending on the relative expansion of two different substances. The case of stretched rubber is well known in that it shows a positive expansion coefficient at right angles but an apparently negative expansion parallel to the direction of stretch. James and Guth[22] have presented a critical interpretation of this behavior. F. Horst Muller[23] has apparently observed a case with rubber where the expansion coefficient was 0.7×10^{-4} parallel, and 1.4–2.3×10^{-4} perpendicular to the chains, but details of his experiments are not available to us at the present time. Muller,[15] from his x-ray studies of *n*-paraffins, has found that both the linear expansion and the compressibility are greater at right angles than parallel to the long axis of the molecules. He concludes that at the transition temperature rotation of the molecule as a whole about its own axis commences. He presents dielectric polarization measurements on long chain ketones to confirm this interpretation. In the case of high polymers rotation of segments rather than of the entire molecule presumably comes into play.

In addition to the non-isotropic thermal expansion of oriented polymers, it was also predicted that $\Delta\beta$ would be affected by orientation. Table II shows the magnitude of this effect for both saran and polystyrene. In each case it is seen that $\Delta\beta$ is smaller for oriented polymers. One might expect that β_2 would be roughly 50 percent greater than β_1 for oriented materials, since expansion in a third direction makes its appearance at *Tm*. This, however, is only approximately true. In fact, $\Delta\beta$ seems to depend on a number of factors, some of which we plan to discuss elsewhere.

It thus seems possible to trace the break in the volume temperature curve of a high polymer—the so-called second-order transition—to two distinct mechanisms of thermal expansion, both of which are acting above *Tm*, but only one being operative below. The sideward motion of chain segments occurs instantly as compared to the time scale of the experiment, whereas the lengthwise motion of chain segments is almost neglible until the critical temperature, *Tm*, is attained. It thus reduces to a problem in the viscous flow of polymer molecules with the consequent necessity of rotation about carbon-carbon bonds, according to the theory of Kauzmann and Eyring.[24] The other related aspects of the second-order transition, such as the discontinuity in the specific heat *versus* temperature curve, the increase in dielectric constant and power factor above a certain temperature, and the onset of long range elasticity are likely understandable in terms of the free rotation about carbon-carbon bonds which becomes possible at *Tm*. This does not mean that segment rotation is non-existent below *Tm*, but rather that as *Tm* is approached the proportion

TABLE II. Effect of orientation on $\Delta\beta$.

Material	$\Delta\beta$
Saran B, oriented	3.2×10^{-4}
Saran B, unoriented	4.0
Polystyrene, oriented	4.8
Polystyrene, unoriented	6.6

[21] R. Vieweg and W. Schneider, Kunststoffe **32**, 295 (1942).
[22] H. M. James and E. Guth, J. Chem. Phys. **11**, 455 (1943).
[23] F. Horst Muller, *Chem. u. Tech. der Kunststoffe* (Edwards Brothers, Ann Arbor, Michigan, 1944), edited by R. Houwink, Vol. I, p. 262.
[24] W. Kauzmann and H. Eyring, J. Am. Chem. Soc. **62**, 3113 (1940).

of bonds or segments which are rotating becomes great enough to accommodate the experimental effect being observed. This is particularly evident in dielectric dispersion measurements where the temperature must be raised with increasing frequency of applied voltage to obtain a given dielectric loss; that is, to obtain a given proportion of polar groups which can rotate at that frequency.

Returning to the thermal expansion problem which will now be examined in more detail, the following definition is proposed:

The second-order transition temperature associated with thermal expansion of polymers is the lowest temperature at which an observable amount of viscous flow can occur under the thermal expansion forces acting on the molecules, and within the time limits of the experimental technique employed.

Actually, for most polymers the transition occurs not at a sharp point but over a finite temperature interval (depending on the rate of heating and other factors) within which the chains are approaching but have not yet attained equilibrium. Aside from some work by Jenckel and Ueberreiter[25] demonstrating the effect of polymolecularity and plasticizers on the sharpness of the transition, there is little available in the way of experimental data. Wood[26] has studied the thermal expansion of stark rubber near *Tm* by a precise interferometric method which showed that the transition was smooth and continuous over a 4° interval. In practice *Tm* is found by extrapolating the linear portions of the volume-temperature curve on either side of the transition interval to a point of intersection.

The definition of *Tm* just outlined is immediately useful because it involves three variables, a thermal *force*, acting against a viscous *resistance* resulting in a certain *rate* of deformation. How then does *Tm* depend on these variables? Different rates of heating, change in the melt viscosity of the polymer by variation in molecular weight or by the use of plasticizer, or finally, augmenting the thermal forces with externally applied loads should all change *Tm* in a predictable direction.

Tammann[27] and Jenckel[11] have pointed out that transition temperatures in glasses correspond to a viscosity of 10^{13} poises, while Wiley[28] has made some calculations on polymers which lead to this same value. It will, therefore, be assumed that *Tm* represents an isoviscous state of the polymer for any set of experiments in which an apparatus of fixed sensitivity is operated at a constant rate of heating. *Tm* will then increase or decrease in response to any variable factor in such a way as to occur always at a fixed value of polymer viscosity. It is not necessary to know the viscosity, but only to assume its constancy.

It is convenient then to treat this problem quantitatively in terms of the melt viscosity equation for polymers either in the form developed experimentally by Flory,[29] or in the more detailed theoretical expression given by Kauzmann and Eyring.[24] Flory was able to show that the melt viscosity, η, of linear polyesters at temperatures at least 50°C above their melting points is given by

$$\ln \eta = D + BZ_w^{\frac{1}{2}} + C/RT, \qquad (4)$$

where B, C, and D are constants, T is the absolute temperature, R is the gas constant, and Z_w is the weight-average chain length of the polymer. Kauzmann and Eyring have provided explicit physical interpretation for the constants in the above equation. Thus it is possible to write

$$\eta = \{(Qf/RT) \exp(-\Delta S/R) \times \exp(\Delta H/RT)\}/\sinh(Pf/RT), \qquad (5)$$

where P and Q are constants, f is the shearing force per unit area, ΔH is the energy of activation for a segment, and ΔS, the entropy term, is dependent on the square root of the weight average chain length. The equation in this form shows the functional dependence of viscosity on shearing force. When $f \ll RT/P$, then $\sinh(Pf/RT) \approx Pf/RT$, and viscosity is independent of shearing force. If, however, $f \gg RT/P$, then

$$\sinh(Pf/RT) \approx (1/2) \exp(Pf/RT).$$

Using this latter relationship, and assuming that the transition temperature represents an isoviscous state, we find that

$$\ln f + \Delta H/RT_m + Pf/RT_m = \text{constant}. \qquad (6)$$

Neglecting $\ln f$ as compared to f, we see that

$$T_m = A' - B'f, \qquad (7)$$

where A' and B' are constants. Thus the transition temperature, *Tm*, should decrease linearly as the applied shearing force f is increased. Between the two extremes (independence of viscosity, and hence *Tm*, on shearing force, and the above

[25] E. Jenckel and K. Ueberreiter, Zeits. f. physik. Chemie A182, 361 (1938).
[26] L. A. Wood, N. Bekkedahl, and C. G. Peters, J. Research Nat. Bur. Stand. 23, 571 (1939).
[27] G. Tamman, *Der Glaszustand* (Leipzig, 1933).
[28] F. Wiley, Ind. Eng. Chem. 34, 1052 (1942).
[29] P. J. Flory, J. Am. Chem. Soc. 62, 1057 (1940).

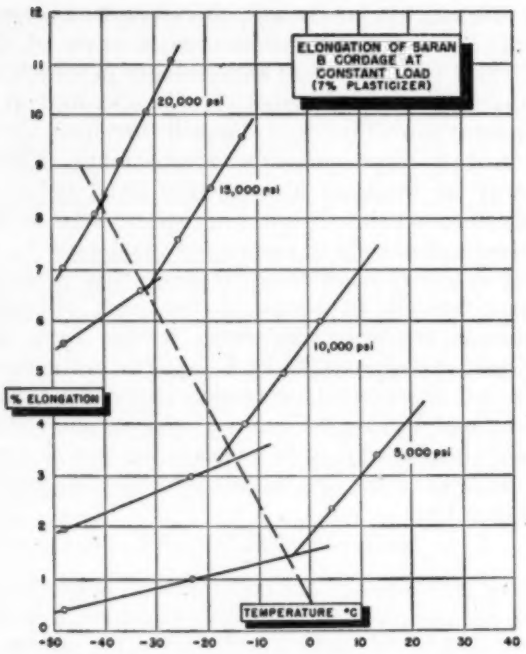


FIG. 4. Length of saran cordage as a function of temperature at several constant loads.

equation) the relationship will be complex, since higher order terms in the expansion of the hyperbolic sine will be involved.

Effect of Applied Forces on *Tm*

As an example of the effect of an applied force on the second-order transition temperature, the data of Lowry and Williams[30] shown in Fig. 4 can be cited. Samples of saran *B* (vinylidene chloride-vinyl chloride copolymer) cordage 0.020″ diameter and containing 7 percent of plasticizer were cooled to −48°C and the indicated loads applied. The samples were then allowed to warm slowly, the length being followed by an automatic recording device. The elongation *versus* temperature curve showed characteristic breaks whose position along the temperature axis varied approximately linearly with the applied stress, much as Eq. (7) might predict. Wood[31] has pointed out that an elastic deformation was undoubtedly present, and could have been determined by recycling back and forth over the temperature scale. Moreover, there may have been some complication with increasing crystallinity and enrichment of plasticizer in the amorphous phase, as discussed in the first paper in this series. A second group of samples containing a different plasticizer behaved in similar manner except that the transition temperature became constant at −34°C when the load reached 15,000 p.s.i. It will be noted in Fig. 4 that the initial elongations at −48°C are not uniformly proportional to the applied load. This was apparently due to an error in adjusting the zero point of the automatic recorder.

Horst Muller[32] has reported similar data on the elongation of polyvinyl chloride foils under load, in which case there should be no complication from crystallization, although there was very likely an elastic component present. The range of loads covered, from 20 to 2000 grams per sq. mm, apparently lies intermediate between the two extremes previously mentioned (viscosity independent of shearing force, and viscosity exponentially dependent on shearing force). At any rate, the relationship between force and transition temperature is not linear. Instead, a plot of the logarithm of applied force against the reciprocal of the transition temperature, as shown in Fig. 5, results in quite a good straight line. It will

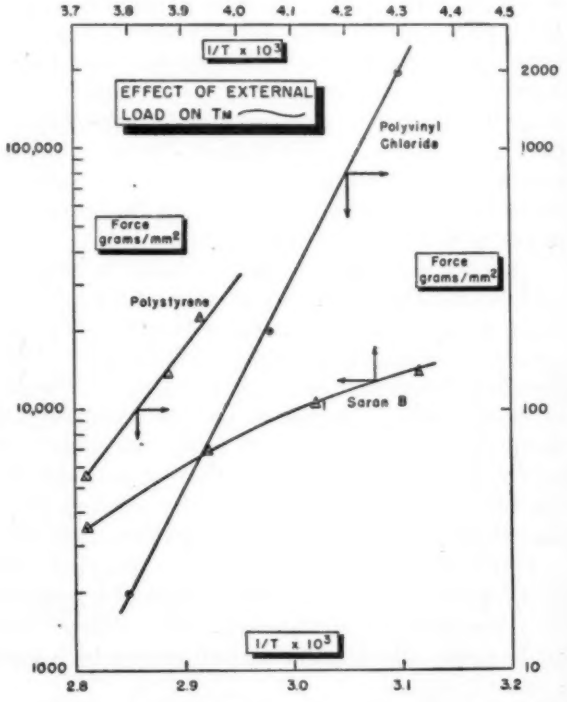


FIG. 5. Variation of the reciprocal transition temperature with external load for polystyrene, polyvinyl chloride, and saran B.

[30] R. D. Lowry and J. L. Williams, The Dow Chemical Company, unpublished data.

[31] L. A. Wood, National Bureau of Standards, private communication.

[32] F. Horst Muller, Kolloid. Zeits. 95, 138 (1941).

be noted that the saran data do not give a straight line when plotted in this manner. The third curve shown in Fig. 5 is from data on the expansion of oriented polystyrene, also obtained by Horst Muller. Here an internal force was frozen into the specimen and produced a contraction at the transition temperature. The greater the orientation, the lower the temperature at which this contraction began. The magnitude of these internal forces was estimated by the specific double refraction exhibited by the oriented samples, using data given by Muller.

It was anticipated that extrapolation to room temperature of the $\log f$ *vs.* $1/T$ curve of Fig. 5 for polyvinyl chloride might give a force corresponding roughly to the yield point of this material. Thus, the yield point might be thought of as the force required to reduce *Tm* to room temperature. However, the extrapolation indicated a yield value of 33,000 p.s.i., which is much too high. Turner,[33] in studying the rate of elongation of polyvinylchloride-acetate films at room temperature, found an exponential increase in rate once a certain critical load was exceeded. This had the characteristics of a transition phenomenon, which suggested, in line with the above ideas, that even at room temperature enough segmental action occurred to allow viscous flow, if only a sufficient force were applied.

Effect of Polymer Melt Viscosity on *Tm*

In our treatment of second-order transitions as a problem in viscous flow, the melt viscosity of the polymer would appear to be the most important property to study. We have already seen how it depends on shearing force and have used this to explain the influence of external loads on the transition temperature. The melt viscosity also depends on the molecular weight of the polymer, as well as on the presence of plasticizers.

TABLE III. Effect of molecular weight on the second-order transition temperature of polystyrene (Jenckel and Ueberreiter).

Molecular weight	*Tm*
6,099	53.5°C
33,000	92.5
58,700	103.0
63,000	95.25
190,000	102.0
255,000	108.5
484,000	105.75

[33] Quoted by E. G. Couzens and W. G. Wearmouth, J. Soc. Chem. Ind. (Trans.) **61**, 69 (1942).

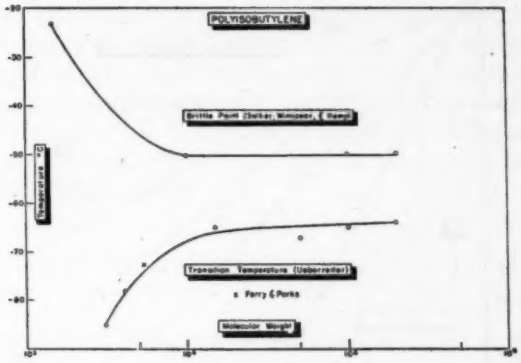


FIG. 6. Variation of brittle point and second-order transition temperatures of polyisobutylene with molecular weight.

Hence, *Tm* should be a function both of molecular weight and plasticizer content. The melt viscosity also depends on temperature; but, under the assumption of an isoviscous state, temperature becomes the dependent variable.

If the molecular weight of a polymer is low, it should reach a melt viscosity of 10^{13} poises (or some other viscosity corresponding to the time scale of the particular experiment) at a lower temperature than would a high molecular polymer of the same type, and *Tm* should accordingly be lower. Jenckel and Ueberreiter[25] have shown that this is indeed the case with polystyrene and with polyisobutylene. Their data is given in Table III and in Fig. 6 respectively. The increase in *Tm* with increasing chain length is rather pronounced in the region of low molecular weights until *Tm* finally approaches a constant value. This implies that above a certain molecular weight the entire molecule is not engaged in a cooperative movement in a thermal expansion experiment. Ueberreiter[34] has, therefore, used this method to fix the length of the kinetically active segment length at 100 monomer units for polystyrene.

For the region of molecular weights where *Tm* is changing rapidly it would be expected that Eq. (4) expressed in the following form:

$$Z_w^{\frac{1}{2}} = G + H/RT \qquad (8)$$

should describe the experimental results. This assumes $\log \eta =$ a constant. There are not available sufficient data to check this point, except in the qualitative sense that *Tm* does depend on molecular weight in the correct manner. Tuckett,[35] however, has applied Eq. (8) to the softening

[34] K. Ueberreiter, Kolloid. Zeits. **102**, 272 (1943).
[35] R. F. Tuckett, Trans. Faraday Soc. **39**, 158 (1943).

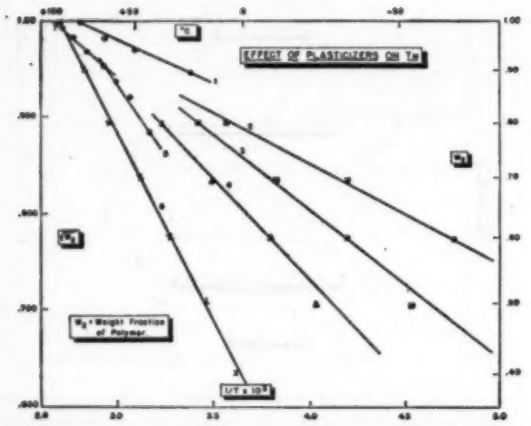


FIG. 7. Lowering of the second-order transition temperature by plasticizers. Curve 1, *Tm* for polystyrene-ethylbenzene; 2, 3, 4, flex brittle temperature for polyvinyl chloride—acetate with different plasticizers; 5, 6, peak in loss factor—temperature curve for polyvinyl chloride plasticized with diphenyl and tricresyl phosphate respectively.

point data of a series of polyvinyl acetate samples covering a wide range of molecular weights. This particular test involved sufficient plastic flow so that in all cases the entire polymer molecule was involved in coordinated motion. He found excellent agreement.

The simplest method of varying the melt viscosity of a high polymer is through the addition of a plasticizer. Since a large body of data exists concerning the role of plasticizers on transition temperatures, electrical properties, and brittle points of polymers, it has been possible to make a rather exhaustive study of the isoviscous concept along these lines. Flory's modification of the melt viscosity equation to the case of concentrated solutions is most useful for this purpose. Flory[36] suggested that if the solvent could be regarded as a low molecular weight polymer of the same chemical species as the high polymer in question then a weight average chain length taken over the solvent and the polymer molecules combined could be substituted in Eq. (4) for Z_w. The equation then assumes the form

$$\ln \eta = D + KW_2^{\frac{1}{2}}, \qquad (9)$$

where D and K are constants and W_2 is the weight fraction of polymer in the mixture. Flory tested this relationship for polydecamethylene adipate polyesters dissolved in diethyl succinate and found it adequate when more than 10 percent of polymer was present. Baker, Fuller, and Heiss[37] had previously pointed out that melt viscosity studies of concentrated solutions might provide a key to plasticizer action.

We must also know how the melt viscosity of plasticized systems depends on temperature. If we assume a temperature dependence similar to that of Eq. (4) for pure polymers, then

$$\ln \eta = D' + KW_2^{\frac{1}{2}} + E/RT, \qquad (10)$$

where E is the activation energy of viscous flow. Postulating the isoviscous state, that is $\ln \eta$ constant, as a criterion for the second-order transition temperature, we obtain the result:

$$W_2^{\frac{1}{2}} = P - E/KRTm. \qquad (11)$$

This equation predicts a linear relationship between reciprocal transition temperature and square root of the weight fraction of polymer in the plasticized mass. Figure 7, where $W_2^{\frac{1}{2}}$ has been plotted against reciprocal absolute temperature for six different cases, indicates a rather satisfactory agreement with Eq. (11). Curve 1 is Ueberreiter's transition data for polystyrene—ethylbenzene mixtures based on volume temperature data.[38] Curves 2, 3, and 4 are Clash and Berg's flex brittle temperature data on vinyl chloride—acetate polymers plasticized respectively with triethylene glycol di-2-ethyl hexoate, di-2-ethyl hexyl phthalate, and tricresyl phosphate.[39] Since this test involved a rather large angular displacement in a short time, the results are complicated by the presence of elastic deformation in addition to viscous flow. That these plasticizers have different molecular weights, shapes, and chemical constitutions likely accounts for the fact that these three curves have different slopes and do not extrapolate to the same value of *Tm* for pure polymer ($W_2^{\frac{1}{2}} = 1.0$). Clash and Rynkiewicz[40] have considered these same data in a different light, calculating what they termed a molal depression of the flex brittle temperature by plasticizers. In this way they were able to compare the efficiency of several plasticizers in lowering the brittle temperature. Curve 5 represents Fuoss's data[41] on the temperature of the peak in the loss factor *versus* temperature curves for varying amounts of diphenyl as plasticizer in polyvinyl chloride. The break in the curve at roughly 10 percent plasticizer corresponds to a similar break observed by Fuoss for

[36] P. J. Flory, J. Phys. Chem. **46**, 870 (1942).
[37] W. O. Baker, C. S. Fuller, and J. H. Heiss, J. Am. Chem. Soc. **63**, 3316 (1941).
[38] K. Ueberreiter, Angew. Chemie **53**, 247 (1940).
[39] R. F. Clash, Jr. and R. M. Berg, Ind. Eng. Chem. **34**, 1218 (1942).
[40] R. F. Clash, Jr. and L. M. Rynkiewicz, Ind. Eng. Chem. **36**, 279 (1944).
[41] R. Fuoss, J. Am. Chem. Soc. **63**, 378 (1941).

density-composition curves. Curve 6 portrays similar electrical data by Davies, Miller, and Busse[42] for polyvinyl chloride plasticized with tricresyl phosphate. This data for 60 cycles is further confirmed by 1000-cycle and 15-cycle data (not shown here) on the same samples. All three curves exhibit excellent linearity and are almost parallel, the 1000-cycle line having a slightly greater slope.

Wurstlin[43] has studied both second-order transitions (from volume-temperature curves) and 50-cycle loss-factors on polyvinyl chloride, after-chlorinated polyvinyl chloride, and polyvinyl acetate, all plasticized with up to 30 percent of tricresyl phosphate. His data give plots similar to those of Fig. 7, the straight lines for the electrical data having a steeper slope than those for Tm. It is recognized that dielectric measurements do not involve viscous flow of polymer chains in the same sense that thermal expansion does. However, there is a close relationship between electrical phenomena and transition effects which makes this discussion pertinent, as was pointed out recently by Tuckett.[44] Finally, mention should be made of Ueberreiter's studies[45] of thermal expansion for ethyl cellulose and nitrocellulose plasticized with tricresyl phosphate. While the entire concentration range from zero to 100 percent plasticizer was studied, linear plots of $W_2^{\frac{1}{2}}$ *vs.* $1/Tm$ were obtained on ethyl cellulose only between 20 and 60 percent of plasticizer; on nitrocellulose from 10 to 40 percent of plasticizer.

These many excellent examples of a linear relationship between $W_2^{\frac{1}{2}}$ and $1/Tm$ would appear at first glance to confirm the role of viscosity, and particularly the concept of the isoviscous state, in second-order transitions. Actually, however, only one of the systems, polystyrene-ethylbenzene, meets Flory's criterion of a plasticizer structurally similar to the base unit in the polymer chain. Secondly, the implicit assumption that the activation energy, E, in Eqs. (10) and (11) is independent of plasticizer content, does not hold true. Kauzmann[46] has already shown that the activation energy for the dielectric dispersion process drops off rather markedly with increasing plasticizer content, while the entropy of activation behaves in similar fashion.

We have recently examined a considerable body of viscosity data obtained as a function of temperature and concentration for the system polystyrene—isopropyl benzene. When this data was analyzed in terms of Eq. (10) it was found that neither K nor E were constant. Instead, a more suitable form for the viscosity equation could be written as

$$\ln \eta = A' + B'Z_w^{\frac{1}{2}} + W_2^{\frac{1}{2}}\left(C' - D'Z_w^{\frac{1}{2}} + \frac{E'}{RT}\right), \quad (12)$$

where A', B', C', and D' are constants, E' is the activation energy divided by $W_2^{\frac{1}{2}}$, and Z_w is the weight average chain length. When Z_w is constant, this equation simplifies to

$$\ln \eta = A'' + W_2^{\frac{1}{2}}\left(C'' + \frac{E'}{RT}\right). \quad (12a)$$

When this equation is subjected to the condition $\ln \eta = \text{constant}$, then

$$1/W_2^{\frac{1}{2}} \propto 1/Tm. \quad (13)$$

Unfortunately, this equation does not give linear plots for the data presented in Fig. 7. The net result is that Eq. (11) must be considered for the present as an empirical relationship. Only for the special case where the ratio E/K in Eq. (11) is a constant independent of temperature and plasticizer content can we still assume that the straight line plots of Fig. 7 imply an isoviscous state. E/K is certainly not constant for polystyrene—isopropyl benzene.

Time Effects in Connection with Second-Order Transition Phenomena

It was previously mentioned that Jenckel's[11] experiments in prolonging the heating period of glassy selenium from hours to several days reduced Tm from 30 to 23°C. While this seems like a negligible effect, yet Jenckel points out that the viscosity has increased from 10^{13} to 10^{16} poises in this temperature interval. Any further lowering of the transition temperature would quickly lead to geological time scales. Time effects can be treated by using Eyring's[47] absolute reaction rate equations dealing with viscous flow. He shows, for example, that the rate dx/dt at which two layers of molecules are displaced relative to each other by an applied force f is:

$$dx/dt = A \exp(\Delta F/RT) \sinh(f\delta/T), \quad (14)$$

[42] J. M. Davies, R. F. Miller, and W. F. Busse, J. Am. Chem. Soc. **63**, 361 (1941).
[43] F. Wurstlin, Kolloid. Zeits. **105**, 15 (1943).
[44] R. F. Tuckett, Trans. Faraday Soc. **40**, 448 (1944).
[45] K. Ueberreiter, Zeits. f. physik. Chemie **B48**, 197 (1941).
[46] W. Kauzmann, Rev. Mod. Phys. **14**, 12 (1942).
[47] J. F. Kincaid, H. Eyring, and A. E. Stearn, Chem. Rev. **28**, 301 (1941).

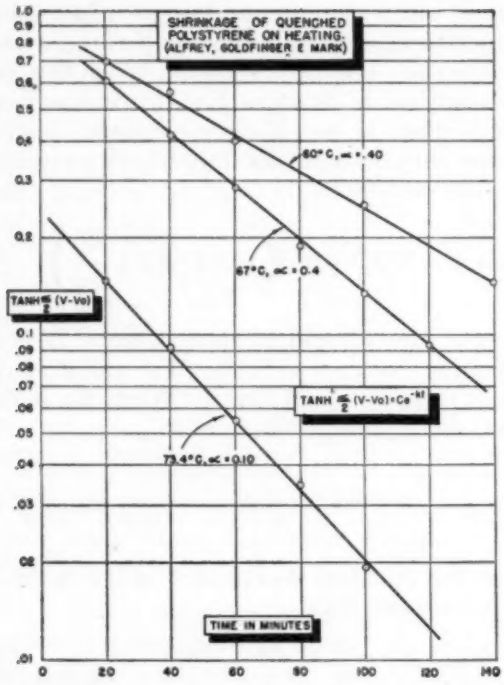


FIG. 8. Shrinkage with time of polystyrene which had been quenched from 95 to 20°C and then held at the indicated temperatures.

where A and δ are constants, and ΔF is the free energy of activation. This equation suggests that if the time scale (that is dx/dt) of the experiment is altered, the reciprocal transition temperature will vary logarithmically, neglecting the presence of T in the hyperbolic sine. There does not appear to be available sufficient thermal expansion data which would permit a precise checking of this point.

There are, however, several sets of experiments such as those of Jenckel, where the volume of a sample is followed with time after a sudden change in temperature. Jenckel found empirically for selenium glass that the volume decreased in proportion to the square root of the time.[11] Alfrey, Goldfinger, and Mark[10] have carried out some instructive experiments on polystyrene as follows: Samples were heated in a dilatometer to 95°C and then suddenly quenched to 20°C. The samples had a greater volume following this quench than they originally possessed at 20°C. On heating after quenching, the expansion coefficient was normal until temperatures around 40°C were reached, whereupon the specimens would contract in an effort to approach their equilibrium volumes. Change of volume with time was studied for quenched samples subsequently heated to 60°, 67.2°, and 73.4°C. The curves appeared to be exponential in character, but actually were not according to the authors.

Treating this contraction as a problem in isothermal viscous flow, it can be assumed that the force causing the contraction is proportional to $V-Vo$, where V is the volume at any time t and Vo is the equilibrium volume after an infinite time at constant temperature. Substituting $V-Vo$ for f in Eq. (14) gives as the rate of change of volume with time:

$$dV/dt = B \sinh [\alpha(V-Vo)], \qquad (15)$$

where α is a constant of proportionality, and where dV, the infinitesimal change in volume, is assumed to be a measure of the relative displacement of chain segments, dx. Integration leads to an expression of the form:

$$\tanh [\alpha(V-Vo)/2] = Ce^{-kt}. \qquad (16)$$

Their data are plotted on semi-logarithmic graph paper in Fig. 8 where it is seen that a rather satisfactory fit is obtained. Jenckel's data on selenium behave equally well on this type of plot. The parameter α contains the so-called viscous volume or the volume of the element which acts as a unit in viscous flow. The fit of the straight lines is relatively insensitive to α, although α shows a substantial drop as the transition temperature (75°C) is approached. The curves of Fig. 8 were fitted by assuming several values of α until a good straight line was obtained. The constant k increased in the order 2.01, 3.52, and 4.15×10^{-4} (for time in seconds) as the temperature changed from 60 to 67.2 to 73.4°C. However, k drifted slightly depending on the numerical values of the time used to compute it. Incidentally, Eq. (16) is of the type used by Tobolsky and Eyring[48] to account for creep under constant stress as well as stress relaxation at constant elongation.

THE BRITTLE POINT OF HIGH POLYMERS

Reference has been made several times to the so-called brittle point of high polymers as another manifestation of second-order transitions. Early writers on the subject considered the brittle temperature, T_b, to be identical with Tm. However, examination of the available experimental data now shows that while T_b and Tm are almost equal for high molecular weight polymers, yet they diverge markedly in the region of low molecular weights. Analysis of the brittle point

[48] A. Tobolsky and H. Eyring, J. Chem. Phys. 11, 125 (1943).

test indicates that such behavior is to be expected. Figure 6, previously cited as showing variation of Tm with molecular weight for polyisobutylene, also contains a plot of brittle temperature on a different set of samples, as reported by Selker, Winspear, and Kemp.[49] While there are only four points on this brittle temperature curve, yet these same authors report similar trends for polyethylene and natural rubber. We have reported elsewhere[50] an identical type of behavior for polyvinylidene chloride brittle points.

As the name implies, the brittle point is the temperature at which a specimen becomes brittle according to some specified test, which may involve a quick or slow bend through a large or small angle, the sudden application of load as in the shatter test, or the measurement of hardness, Young's modulus, or some similar property. Most of the tests in use require that a sample undergo a fixed deformation in a given length of time. If it can meet this condition it passes the test, if not, it breaks and fails. There are three mechanisms by which this deformation can occur:

1. An ordinary elastic deformation obeying Hooke's law, having a modulus of around 10^{11} dynes per sq. cm, and almost instantaneous. This will usually make a rather negligible contribution to the total deformation, except in the case of very thin specimens.

2. Highly elastic deformation resulting from the uncoiling of polymer chains. This mechanism is characterized by a relaxation time, or distribution of relaxation times, and is generally written in the form

$$D_{he}=D_{\infty}[1-\exp(-t/\tau)], \qquad (17)$$

where D_{∞} is the deformation after infinite time, τ is the relaxation time, and t is the elapsed time. τ shows an exponential decrease with temperature and also depends on the molecular weight.[16]

3. Viscous flow. This is a relatively slow process, with the viscosity being given by Flory's melt viscosity equation, as seen in Eq. (4).

In Fig. 9 we see the contribution of these three mechanisms to an observed deformation at different temperatures. The ordinary elastic deformation is almost instantaneous, followed by the highly elastic deformation. At longer times viscous flow becomes appreciable and eventually the deformation curve becomes linear with time. With increasing temperature the curve tilts toward higher deformation. The brittle point

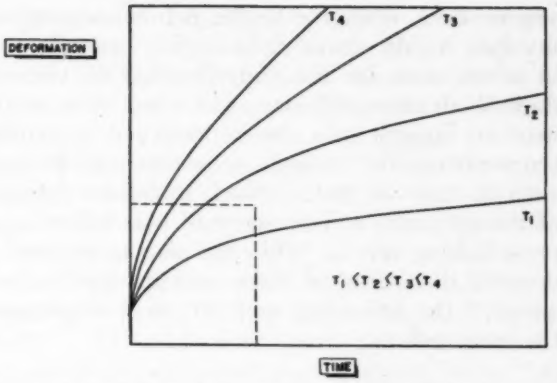


FIG. 9. Deformation—time curves of a polymer at several temperatures, all of which are above the second-order transition temperature. The dotted lines represent the deformation required by the brittle-point test, and the time available for this deformation.

test requires a fixed deformation in a fixed time interval, as indicated by the dotted lines on this figure, and thus we arrive at the following definition of the brittle point: *The brittle temperature is that temperature at which the time interval required by the sample to undergo the necessary deformation is just equal to the time interval of the test.* On our figure this means that at this temperature the deformation-time curve passes through the point of intersection of the two dotted lines. At higher temperatures the time is more than sufficient and the sample deforms without breaking; at lower temperatures the sample cannot deform fast enough and it breaks.

The brittle-point test involves a relatively large and relatively rapid deformation of the sample as compared to a volume-temperature determination of the transition temperature. Tuckett has shown that Tm is a measure of the temperature, or temperature region, in which high elasticity and viscous flow can develop in a polymer. Tm might then be considered as a lower limit for T_b, in the sense that the faster the brittle-point test, and the greater the deformation required by the test, the higher will T_b lie with respect to Tm. Conversely, when the deformation and/or the rate of flexing is small, T_b will approach Tm. Since some of the early brittle-point tests, especially that due to Russel,[51] used a relatively slow deformation, it is readily seen how Tm and T_b came to be identified as meaning the same thing. These two tests might be differentiated by saying that second-order transition temperatures obtained from volume-temperature curves represent a problem in

[49] M. L. Selker, G. G. Winspear, and A. R. Kemp, Ind. Eng. Chem. **34**, 157 (1942).

[50] R. F. Boyer and R. S. Spencer, *Advances in Colloid Science* (Interscience Publishers, Inc., New York City), Vol. II, in press.

[51] J. J. Russel, Ind. Eng. Chem. **32**, 509 (1940).

viscous flow, while the brittle-point temperature involves highly elastic deformation, complicated in some cases by the superposition of viscous flow. With these differences in mind, it is pertinent to inquire into the dependence of brittle temperature on various experimental factors such as type of test method, molecular weight of the polymer, and presence of plasticizers and cross-linking agents. While we plan to present a detailed discussion of these various effects elsewhere,[50] the following summary will emphasize the main points.

1. Effect of Deformation and Speed

Kemp, Malm, and Winspear[52] have studied the effect of amount of deformation and rate of deformation on the brittle point of plasticized polyvinyl chloride. Some of their findings, presented in Table IV, emphasize quite well the fact that T_b is dependent on these variables in the qualitative manner required by our definition. A further extension of this idea in regard to sample thickness has been found by us for saran.[50] Here a thicker sample would presumably experience a larger deformation. The brittle temperature increased from roughly $-50°C$ to $-10°C$ as specimen thickness increased from 5 to 70 mils. It appeared from the data that ordinary elasticity may have provided most of the deformation when the sample thickness was below 10 mils.

TABLE IV. Brittle point of plasticized polyvinyl chloride (Kemp, Malm, and Winspear).

Rate of deformation	Bending radius	Brittle temperature
150 r.p.m.	Right angle	−25.5°C
	2.54 cm	−32.5
	3.81 cm	−46.5
75 r.p.m.	Right angle	−30.0
	2.54 cm	−38.5
	3.81 cm	−51.5

2. Effect of Molecular Weight

We have already seen in Fig. 6 that the brittle point of polyisobutylene drops rapidly with increasing molecular weight and then approaches a constant value at higher molecular weights where commercial polymers are likely to occur. Long hydrocarbon chain polymers above their second-order transition temperature can develop a total highly elastic deformation of several hundred percent, but the brittle point test requires only a small fraction of this maximum possible deformation. Hence it might be expected that above a certain chain length T_b would no longer depend on molecular weight. However, with progressively decreasing chain length, the property of long range elasticity would gradually disappear. Highly elastic deformation could no longer supply the deformation required by the brittle point test, and only viscous flow would remain to keep the sample from breaking. The extremely high viscosity of a polymer (about 10^{13} poises at Tm) means that viscous flow is very sluggish unless the temperature is raised considerably above Tm. The increase in T_b at very low molecular weights probably represents the decreasing contribution of elastic deformation and the increasing contribution of viscous flow to the total deformation needed by the brittle-point test. Eventually, at some sufficiently low molecular weight, the polymer might revert to a liquid whose brittle point would decrease. It should be possible to make a rigorous quantitative treatment of the effect of molecular weight on brittle point. However, there is not available sufficient information concerning influence of molecular weight on relaxation times or on the numerical constants appropriate for the melt viscosity equation. An approximate treatment of this problem was carried out, in which the fixed deformation required by a brittle-point test was set equal to the sum of the highly elastic and viscous deformations for a fixed time interval. D_∞ in Eq. (17) was assumed proportional to molecular weight. This treatment suggested that the reciprocal brittle temperature should vary linearly with the square root of the molecular weight. Our data on the brittle points of polyvinylidene chloride followed this prediction reasonably well.[50]

3. Effect of Plasticizer on Brittle Point

It is a well-recognized fact that the presence of a plasticizer lowers the brittle point of a polymer to an extent depending on the amount of plasticizer present.[39, 51] We have previously seen how plasticizer lowers the second-order transition temperature according to the empirical law illustrated in Fig. 7. Since Tm marks the temperature region in which high elasticity develops, and hence represents a lower limit for T_b, it might be expected that T_b would decrease with plasticizer content in about the same manner as does Tm. Indeed the flex brittle temperature

[52] A. R. Kemp, F. S. Malm, and G. G. Winspear, Ind. Eng. Chem. 35, 488 (1943).

curves of Fig. 7 have already indicated that this is approximately true.

In order to check this point more fully, we prepared samples of polyvinyl chloride, 30 mils thick, plasticized with up to 60 percent of tricresyl phosphate. These were measured for brittle point with the apparatus of Selker, Winspear, and Kemp.[49] The resulting data, presented in Table V, has the confusing property

TABLE V. Effect of plasticizer content on the brittle point of polyvinyl chloride.

% Tricresyl Phosphate by weight	Brittle temperature
15	>32°C
20	32
25	28
30	12
35	5
40	− 5
45	−26
50	−27
55	−42
60	−55

of giving three types of linear plots, all about equally good: T_b *vs.* percent plasticizer, $1/T_b$ *vs.* $W_2^{\frac{1}{2}}$, and $1/T_b$ *vs.* $1/W_2^{\frac{1}{2}}$. The latter type of plot, already discussed as Eq. (13), is believed preferable since it extrapolates to roughly 80°C for pure polyvinyl chloride, while the others do not.

4. Effect of Vulcanization

The first effect of small amounts of vulcanization is in suppressing viscous flow[16, 50, 53] although increasing amounts of cross-linking will increase Tm and hence would be expected to increase T_b. Selker, Winspear, and Kemp[49] have made rather extensive studies of T_b as a function of vulcanization. They find that T_b does not increase until about 1 percent of sulphur is reached, and thereafter it increases in regular fashion.

CONCLUSIONS

We have attempted to present a specific mechanism for the thermal expansion process in high polymers. It is based on the idea that the length of a given polymer chain is relatively invariant with temperature; that adjacent chains can move sidewards with respect to each other almost instantaneously at all temperature; that lengthwise relative motion of two chains or chain segments is negligible below a certain critical temperature which we identify with the second-order transition temperature, Tm; that in the neighborhood of Tm this lengthwise flow of chains gives rise to the time effects which have been noted in volume-temperature measurements of polymers; and finally, that above Tm this lengthwise relative motion of chains is rapid enough to follow the temperature rise, thereby accounting for the increased value of thermal expansion above Tm.

We then tried to apply melt viscosity laws to explain various factors which might influence Tm, if the transition point were assumed to represent an isoviscous state. It was found that Tm increased with molecular weight, decreased with applied forces and with increasing plasticizer content, and that the volume near Tm changed with time in qualitative agreement with the ideas of melt viscosity. Various linear laws connecting Tm with external variables were shown graphically, although it still remains to be seen how much of this agreement was fortuitous, or to what extent it represents first approximations to the truth.

Finally, the brittle-point test has been analyzed and shown to differ in several important respects from second-order transitions. This divergence appears most striking for low molecular weight polymers where Tm decreases while T_b increases as the molecular weight is lowered. The brittle-point test represents fast, highly elastic deformation arising from the uncoiling of polymer chains, as compared with the sluggish movement of chain segments during thermal expansion. T_b approaches Tm as a lower limit, but will generally be above it by an amount which depends on the speed of the test, the deformation involved, and the molecular weight of the polymer. Factors which reduce highly elastic deformation, i.e., vulcanization, fillers, and short chains, raise T_b, whereas long chain length or plasticizers, both of which aid highly elastic deformation, lower T_b.

ACKNOWLEDGMENT

This material was presented in part before the Society of Rheology (New York City: October 30, 1943; May 20, 1944) and before the High Polymer Division of the American Physical Society (Evanston, Illinois, November 13, 1943; Rochester, New York, June 24, 1944). We are very much indebted to a number of people who offered comments. We should particularly like to thank R. Simha, L. A. Wood, R. M. Wiley, L. A. Matheson, and T. Alfrey for the privilege of discussing some of these points in greater detail.

[53] D. D. Eley, Trans. Faraday Soc. 38, 299 (1942).

A Method of Measuring the Radiofrequency Resistance of Wires

CHANDLER STEWART, JR.
Headquarters, Army Air Technical Service Command, Radio and Radar Subdivision, Wright Field, Dayton, Ohio
(Received March 6, 1945)

A new and relatively simple method of measuring the resistance of wires at frequencies between about 200 kilocycles and about 40 megacycles is described. By this method, the test specimen is strung down the center of a metal pipe about ten to a hundred feet in length, depending upon the test frequency, to form a transmission line. One end of the line is short circuited and the other connected to a standard commercial Q-meter. Readings are taken from the Q-meter and applied to curves to obtain the desired resistance. This method of measurement appears to be more reliable than the use of coils, and the results agree closely with calculations by skin-effect formulas. For round wires known to consist of homogeneous material, calculated values of radiofrequency resistance appear to be just as reliable and accurate as experimental values, and are much more easily obtained.

INTRODUCTION

THE importance of radiofrequency resistance as a factor in the choice of wire for antennas is particularly great in the case of fixed wire aircraft antennas, where relatively short radiators are operated at low frequencies, resulting in low radiation resistance.[1,2] Consequently, it becomes necessary to determine the r-f resistance of wires contemplated for this use.

The problem which instituted the work to be described involved frequencies used for long range radio communication, from 200 kilocycles to 40 megacycles, and solid, copper clad steel, and flexible braided wire types. Previous work has been done on both the calculation of and measurement of the radiofrequency resistance of wires.[3–10] However, since the formulas for calculation of r-f resistance assume the wires to be homogeneous, they are not applicable to the copper clad steel or flexible braided types used in some radio antennas. Also, experimental work to date has largely been limited to frequencies lower than those generally used for long range radio communication or to difficult methods requiring cumbersome non-standard equipment.

In choosing a method for measuring this resistance, the effect of radiation, proximity, distributed capacitance, and measuring equipment resistance must be considered. Obviously, a very convenient method would be to wind the wire into a solenoid and measure its resistance; however, it was found that the proximity effect and the effect of distributed capacitance of such coils could not be accurately determined. The method appearing to yield the most satisfactory results involved the application of transmission line equations to a coaxial line using the test specimen as a central conductor. This reduces radiation and proximity effects to negligible values, and accurately takes care of distributed capacitance. This method, as developed by the author, will be described.

DEFINITIONS OF TERMS

α = attenuation constant of transmission line in nepers per foot
β = wave-length constant of transmission line in radians per foot
C = distributed capacitance of transmission line in farads per foot
C_0 = capacitance required to resonate the transmission line at the test frequency, in farads
C_1 = capacitance required to resonate reactor No. 1, at frequency f, in farads
C_2 = capacitance required to resonate reactor No. 2, at frequency f, in farads
C_3 = capacitance required to resonate the series combination of test reactors 1 and 2 at frequency f, in farads
C_t = total line capacitance in micromicrofarads $= 10^{12}lC$

f = test frequency in cycles per second
f_{mc} = test frequency in megacycles per second $= 10^{-6}f$

[1] G. L. Haller, "Aircraft Antennas," Proc. I. R. E. **30**, 357 (1942).
[2] G. W. Pierce, *Electric Oscillations and Electric Waves* (McGraw-Hill Book Company, Inc., New York, 1920), Chap. 9.
[3] R. I. Sarbacher and W. A. Edson, *Hyper and Ultrahigh Frequency Engineering* (John Wiley & Sons, Inc., New York, 1944), p. 283, Eq. (7.181).
[4] Sallie Pero Mead, "Wave Propagation Over Parallel Tubular Conductors: The Alternating Current Resistance," Bell Sys. Tech. J. **4**, 331 (1925), Eq. (19).
[5] *Handbook of Chemistry and Physics* (Chemical Rubber Publishing Company, Cleveland, Ohio, 1943), 28th edition, 2399 and 2400.
[6] F. E. Terman, *The Radio Engineers' Handbook* (McGraw-Hill Book Company, Inc., New York, 1943), p. 31 through 37.
[7] John R. Carson, "Wave Propagation Over Parallel Wires: The Proximity Effect," Phil. Mag. (April 1921).
[8] H. B. Dwight, "Proximity Effect in Wires and Thin Tubes," Trans. A. I. E. E. **42**, 850 (1923).
[9] National Bureau of Standards Circular C-74: "Radio Instruments and Measurements," page 299.
[10] A. E. Kennelly, F. A. Laws, and P. H. Pierce, "Experimental Researches on Skin Effect in Conductors," Trans. A. I. E. E. **34**, 1953, September 16, 1915.

G = distributed shunt conductance of transmission line in ohms per foot
$j = (-1)^{\frac{1}{2}}$
l = length of transmission line, in feet
L = distributed inductance of transmission line in henries per foot
n = a positive integer
$Q = X_{sc}/R_{sc}$
Q_m = uncorrected Q-meter reading
Q_1 = Q-meter reading of reactor No. 1 at frequency f
Q_2 = Q-meter reading of reactor No. 2 at frequency f
Q_3 = Q-meter reading of the series combination of reactors 1 and 2, at frequency f
R = series resistance of transmission line in ohms per foot
R_Q = effective residual series resistance of Q-meter at frequency f, in ohms
R_{sc} = input resistance of transmission line at open end, in ohms
R_1 = series resistance at frequency f, of test reactor No. 1, in ohms
R_2 = series resistance at frequency f, of reactor No. 2, in ohms
V = velocity of r-f potential along transmission line, in feet per second = $983.57 \cdot 10^6$ feet per second
W = defined by Eq. (28)
ω = angular velocity of test voltage, in radians per second = $2\pi f$
X_{sc} = input reactance of transmission line at open end, in ohms
Z_0 = absolute magnitude of characteristic impedance of transmission line in ohms
$\dot{Z}_0$ = complex characteristic impedance of transmission line in ohms
$Z_{0c} = R_{0c} + jX_{0c}$
Z_{sc} = input impedance of transmission line at open end, in ohms $R_{in} + jX_{in}$.

DERIVATION OF FORMULAE FOR MEASUREMENT OF RADIOFREQUENCY RESISTANCE OF WIRE

For a line with air dielectric and metal conductors, the following two assumptions are justified from experience:[11]

$$G = 0, \tag{1}$$

and

$$\alpha/\beta \ll 1. \tag{2}$$

From transmission line theory[12] we obtain

$$\beta = \omega(LC)^{\frac{1}{2}}, \tag{3}$$

$$\dot{Z}_0 = \frac{(L/C)^{\frac{1}{2}}}{-R/2\omega L}, \tag{4}$$

and

$$\alpha = R/2Z_0. \tag{5}$$

From (2) and by definition

$$Z_0 = (L/C)^{\frac{1}{2}}, \tag{6}$$

so that dividing (5) by (3) and substituting (6) yields

$$\alpha/\beta = R/2\omega L. \tag{7}$$

[11] See Eq. (21).
[12] W. L. Everitt, "Communication Engineering," page 118, Eqs. (35b), (35d), and (36a).

Substitution of (6) and (7) in (4) yields

$$\dot{Z}_0 = \frac{Z_0}{(-\alpha/\beta)}. \tag{8}$$

On applying assumption (2) this becomes

$$\dot{Z}_0 \cong Z_0[1 - j(\alpha/\beta)]. \tag{9}$$

From transmission line theory[13] we obtain

$$Z_{sc} = \dot{Z}_0 \tanh \gamma l. \tag{10}$$

On substituting the identity $\gamma = \alpha + j\beta$ and converting to real angles, this becomes

$$Z_{sc} = \frac{\dot{Z}_0}{2}\left(\frac{\sinh 2\alpha l + j \sin 2\beta l}{\cosh^2 \alpha l - \sin^2 \beta l}\right). \tag{11}$$

For the condition[14] of

$$\alpha l \ll 1, \tag{12}$$

the assumption

$$\tanh \alpha l \cong \alpha l \tag{13}$$

is justified. Substituting (9) in (11), and making approximations based on (2) and (13) yields

$$Z_{sc} \cong Z_0\left(\alpha l \sec^2 \beta l + \frac{\alpha}{\beta} \tan \beta l + j \tan \beta l\right), \tag{14}$$

from which, by definition

$$R_{sc} = Z_0\alpha\left(l \sec^2 \beta l + \frac{\tan \beta l}{\beta}\right), \tag{15}$$

$$X_{sc} = Z_0 \tan \beta l, \tag{16}$$

$$\frac{X_{sc}}{Q} = R_{sc} = Z_0\alpha\left[l\left(1 + \frac{X_{sc}^2}{Z_0^2}\right) + \frac{X_{sc}}{Z_0\beta}\right]. \tag{17}$$

From which

$$\alpha = \frac{1}{Q[l(Z_0/X_{sc} + X_{sc}/Z_0) + 1/\beta]}. \tag{18}$$

If Q is obtained from a standard Q-meter,

$$Q \geqq 20, \tag{19}$$

so that from (18)

$$\alpha l \leqq 1/20, \tag{20}$$

which supports the statement of (12) and

$$\frac{\alpha}{\beta} = \frac{1}{Q[\beta l(Z_0/X_{sc} + X_{sc}/Z_0) + 1]} \leqq \frac{1}{20}, \tag{21}$$

which supports the statement of (2). From

[13] W. L. Everitt, "Communication Engineering," page 158, Eq. (45).
[14] See Eq. (20).

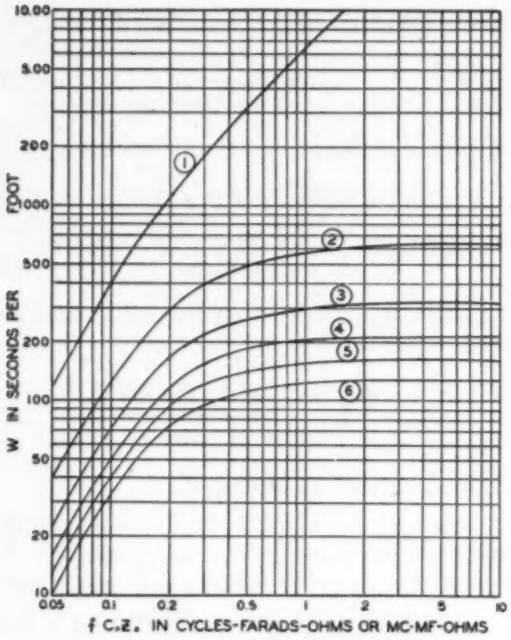


FIG. 1. Plot of $W=2/V[(1+1/(2\pi fC_0Z_0)^2)\operatorname{arccot}2\pi fC_0Z_0+1/2\pi fC_0Z_0]$.

Curve No. 1. $lf_{mc}<246$.
Curve No. 2. $492<lf_{mc}<738$.
Curve No. 3. $984<lf_{mc}<1230$.
Curve No. 4. $1475<lf_{mc}<1721$.
Curve No. 5. $1967<lf_{mc}<2213$.
Curve No. 6. $2459<lf_{mc}<2705$.

transmission line theory[15]

$$\beta=\omega/V=2\pi f/V. \tag{22}$$

If X_{sc} is determined by resonating with a capacitor C_0,

$$X_{sc}=1/2\pi fC_0. \tag{23}$$

Substituting (5), (22), and (23) in (18) yields

$$R=2Z_0\Big/Q\left[l\left(2\pi fC_0Z_0+\frac{1}{2\pi fCZ}\right)+\frac{V}{2\pi f}\right]. \tag{24}$$

This quantity, R, represents the series resistance of the transmission line in ohms per foot, which is the sum of the wire resistance and tube resistance in ohms per foot. The fact that a wire has the same resistance in free space as it has as the central conductor of a coaxial transmission line can be shown by inspection of formulas for the resistance of the central conductor of a coaxial line.[3,4] The latter is not a function of outer conductor inside diameter. Also, these formulas give the same results as do the skin effect formulas for isolated wires.[5,6]

From (16), (22), and (23)

$$\cot\frac{2\pi fl}{V}=2\pi fC_0Z_0. \tag{25}$$

Let

$$W=2\Big/V\left[\left(1+\frac{1}{(2\pi fC_0Z_0)^2}\right)\times\operatorname{arccot}2\pi fC_0Z_0+\frac{1}{2\pi fC_0Z_0}\right]. \tag{26}$$

Then, from (24)—(26),

$$R=W/C_0Q. \tag{27}$$

From inspection of (25) and (26) it is obvious that W has a different set of values for each quadrant in which $2\pi fl/V$ lies. Also it will be noted that the error in W caused by an error in $2\pi fC_0Z_0$ decreases with higher quadrants, provided the quadrant number is known. Since $2\pi fl/V$ and $2\pi fC_0Z_0$ are positive quantities, $2\pi fl/V$ can lie only in positive odd quadrants, so that

$$n\pi<\frac{2\pi fl}{V}<(n+\tfrac{1}{2})\pi, \tag{28}$$

or

$$\frac{nV}{2f}<l<\frac{(n+\frac{1}{2})V}{2f}, \tag{28a}$$

or

$$\frac{nV}{2l}<f<\frac{(n+\frac{1}{2})V}{2l}, \tag{28b}$$

$$\frac{491.8n}{l}<f_{m0}<\frac{491.8n+245.9}{l}. \tag{28c}$$

Equation (26) has been plotted for several quadrants of $2\pi fl/V$ in Fig. 1. From (3), (6), and (22):

$$Z_0=l/VC=1016.7l/C. \tag{29}$$

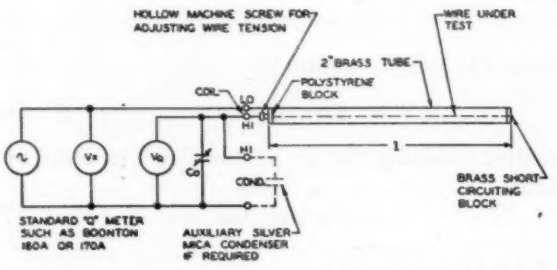


FIG. 2. Arrangement of apparatus for measuring r-f resistance of wires.

[15] W. L. Everitt, "Communication Engineering," page 108, Eq. (23).

From Q-meter theory,[16]

$$R_1+R_Q=1/Q_1\omega C_1, \tag{30}$$

$$R_2+R_Q=1/Q_2\omega C_2, \tag{31}$$

and

$$R_1+R_2+R_Q=1/Q_3\omega C_3. \tag{32}$$

Solving (30)—(32) for R_Q, we obtain

$$R_Q=\frac{1}{\omega}\left[\frac{1}{Q_1C_1}+\frac{1}{Q_2C_2}-\frac{1}{Q_3C_3}\right]. \tag{33}$$

Applying Q-meter theory,[16] to this measurement,

$$Q=\frac{X_{in}}{\frac{X_{in}}{Q_M}-R_Q}=\frac{Q_m}{(1-2\pi fC_0R_QQ_M)}. \tag{34}$$

In practice it will be found that

$$2\pi fC_0R_QQ_M\ll 1, \tag{35}$$

which justifies the following approximation of (34):

$$Q\cong Q_M(1+2\pi fC_0R_QQ_M). \tag{36}$$

EXPERIMENTAL TECHNIQUE

a. Summary of Method

The wire to be tested is strung down the center of a metal tube from a short circuiting plug to a Q-meter, as shown in Figs. 2 and 3. Readings of Q and C_0 are obtained at each test frequency, f, and inserted in the following formula for resistance, R, of the transmission line in ohms per foot

$$R=W/C_0Q \tag{27}$$

where W=function of f, C_0, and Z_0, as defined by Eq. (26), values of which may be taken from the curves in Fig. 1.

Z_0=line characteristic impedance.

The actual wire resistance is obtained by subtracting the calculated resistance of the metal tube from R.

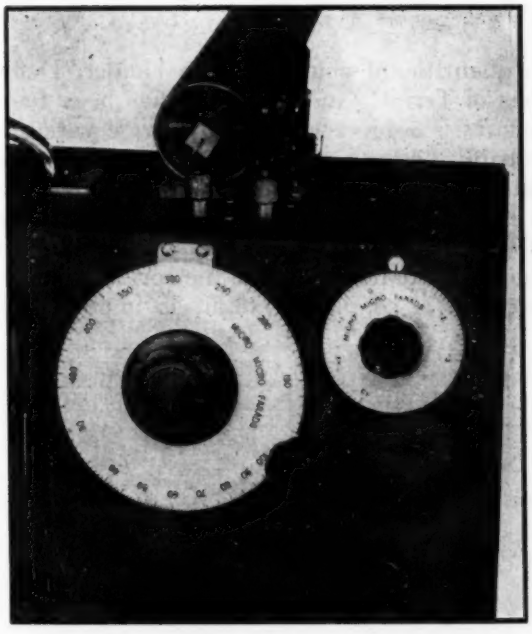


FIG. 3. Connection of coaxial transmission line to Q-meter for measurement of radiofrequency resistance of wires.

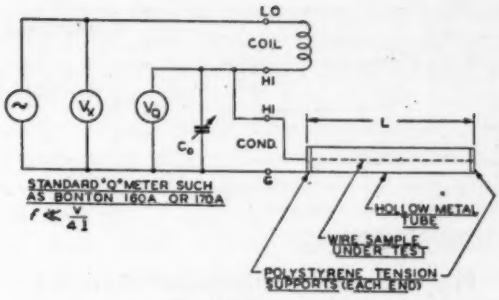


FIG. 4. Arrangement of apparatus for line capacitance determination.

b. Line Lengths and Test Frequencies

The length of line for a given test frequency is limited to ranges which result in inductive reactances, as designated by Eq. (28a). In order that the required resonating capacitance be above the minimum obtainable on the Q-meter, the length is further limited by Eq. (25). Also, in order to minimize the error due to the Q-meter circuit resistance, as can be seen by inspection of Eq. (36), the length and frequency must yield a low C_0Q product.

In practice, it has been found most convenient to use several lengths of line, and to choose test frequencies for each one which will result in low C_0Q products.

c. Resistance of Metal Tube

To obtain the actual resistance of the wire sample, it is of course necessary to subtract from the value of R obtained as above the resistance per foot of the metal tube used. This can be calculated by skin effect formulas.[5,6] By inspection of these formulas it will be seen that this correction, and consequently any error introduced by it, will be materially reduced by employing large diameter tubes, heavily silver

[16] See *Instructions and Manual of Radio Frequency Measurements for Q Meters* (Boonton Radio Corporation, Boonton, New Jersey, June 1941), Eq. (7).

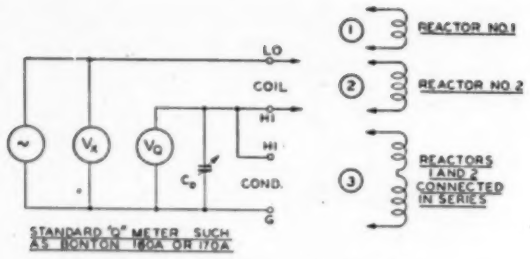


FIG. 5. Circuits for determination of internal resistance of Q-meter.

plated on the inside. The thickness of the plating required will be governed by the lowest test frequency,[17] being thicker for lower frequencies. If frequencies of the order of 1 megacycle and lower are to be used, it may be more convenient to use solid copper tubing instead of plated material.

d. Determination of Line Characteristic Impedance

If the wire sample has a uniform surface, the characteristic impedance, Z_0, of the line may be calculated from[18]

$$Z_0 = 138.5 \log_{10} \frac{r_0}{r_1},$$

where r_0 = inside radius of tube

r_1 = wire radius, in same units.

However, if the wire has a non-uniform or non-cylindrical surface, such as that of either specimen A or specimen B (see Fig. 8), it will be necessary to determine the capacitance of the line experimentally. To do this, the line is open circuited and connected as shown in Fig. 4, and resonated with an inductance by the Q-meter condenser at a low setting. The frequency for this test should be as low as is convenient, and should correspond to a wave-length of at least twenty times the line length. The wire is disconnected from the Q-meter and the condenser re-resonated. The difference in the two condenser settings gives the line capacitance, C_T, which is inserted in the following (previously derived) formula for characteristic impedance Z_0:

$$Z_0 = 1016.7 l / C_T \text{ ohms}, \qquad (29)$$

where l is the length of the line in feet.

[17] F. E. Terman, *Radio Engineers Handbook*, p. 34, Eq. (5).

[18] Keith Henney, *The Radio Engineering Handbook* (McGraw-Hill Book Company, New York, 1941), Third edition, p. 165, Eq. (76).

e. Correction for Q-Meter Internal Resistance

It has been found that the resistance of the Q-meter circuit, R_Q, can introduce a huge error, if operating conditions are not chosen to minimize it or if a correction for this resistance is not applied. From the equation

$$Q = Q_M(1 + 2\pi f C_0 R_Q Q_M), \qquad (36)$$

where

Q = true circuit Q,
Q_M = meter reading,
f = test frequency in cycles per second,
C_0 = resonating capacitance in farads,

it is apparent that this error is least effective when C_0 and Q_M have relatively low values. It has been found in practice that this condition can be met by a discrete choice of frequencies, made while taking data on f, Q_M, and C_0.

In order to determine the magnitude of R_Q, for use in applying a correction to Q_M to obtain Q, the Q and C readings of several inductors at various frequencies are obtained, and the additional readings of Q and C obtained at the same frequencies for pairs of these inductors in series as shown in Fig. 5. R_Q can then be calculated from

$$R_Q = \frac{1}{\omega}\left[\frac{1}{Q_1 C_1} + \frac{1}{Q_2 C_2} - \frac{1}{Q_3 C_3}\right], \qquad (33)$$

the quantities of which are defined under "Definitions of Terms." In practice, it has been found that R_Q is subject to extremely wide variation, probably because of variations in contact area and pressure at the terminals during the tests. In one case it was found to be of the order of 0.075 ohm at one megacycle, increasing somewhat with frequency. Since the correction based upon R_Q is so inaccurate, it is important to

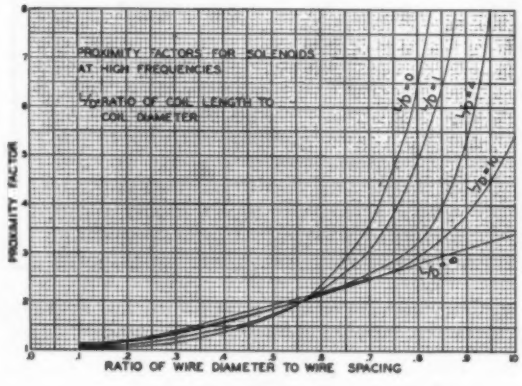

PROXIMITY FACTORS FOR SOLENOIDS AT HIGH FREQUENCIES
RATIO OF COIL LENGTH TO COIL DIAMETER
PROXIMITY FACTOR
RATIO OF WIRE DIAMETER TO WIRE SPACING

FIG. 6.

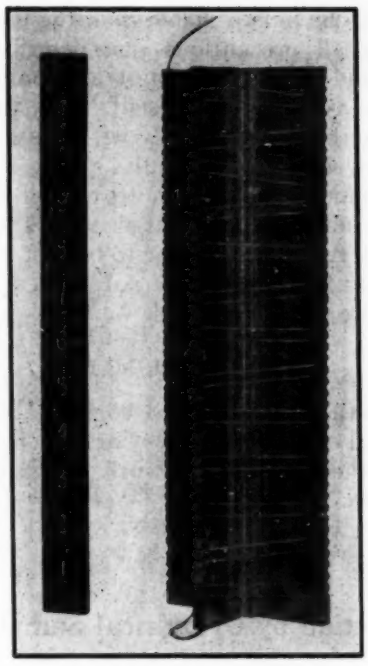


FIG. 7. Low inductance winding for measurement of radiofrequency resistance of wires. Spacing: 0.25 inch; wire diameter: 0.04 inch.

hold C_0Q_M to a relatively low figure, as was just explained.

CHECKS ON THE METHOD

a. Coil Measurements

An approximate formula for the resistance R of straight wire in ohms per foot is

$$R=\frac{1/2\pi fC_0Q_M-(R_Q+R_R)}{Pl}$$

where f, C_0, and Q_M are data taken on a Q-meter of a coil of the wire being tested.

R_Q = Q-meter internal resistance. See (33) and Fig. 5,
R_R = radiation resistance[19] of the coil in ohms,

$$=320\left(\frac{\pi^2}{\lambda^2}NA\right)^2,$$

N = number of turns in coil,
A = area of turn,
λ = wave-length (in consistent units),
l = length of wire in feet,
P = proximity factor[20] (see Fig. 6).

Several measurements were made by this method, as a check on the results of the coaxial transmission line measurements. For reasonable accuracy, it was found necessary to choose frequencies discretely to reduce the effect of R_Q and distributed capacitances, and to use wide turn spacings compared to wire diameter, to reduce the proximity effect. The radiation resistance, R_R, as calculated, was found to be negligible in each case.

In some cases it was necessary to use low-inductance windings, as shown in Fig. 7, in order to reduce Q_M and thereby reduce the error due to R_Q. Typical results obtained at 1 megacycle were as follows:

Specimen A, $\frac{1}{4}''$ spacing, 43 ohms/1000 feet,
Specimen C, $\frac{1}{4}''$ spacing, 34 ohms/1000 feet,
Specimen D, 0.153″ A1 24S, $\frac{1}{2}''$ spacing, 9.6 ohms/1000 feet.

b. Direct Current Resistance

Results of Wheatstone bridge measurements on lengths over ten feet, were as follows:

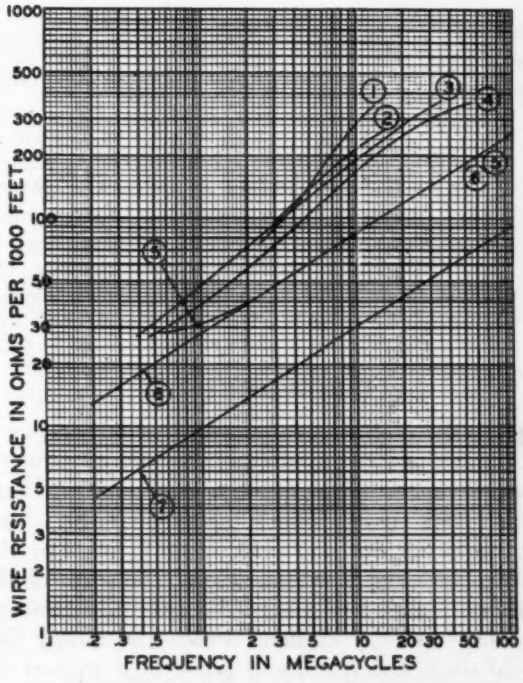


FIG. 8. The radiofrequency resistance of a few typical samples of antenna wire: 1. Bronze braid with cotton inner core, approx. 0.065 inch over-all diameter. 2. Same as No. 1 except tinned bronze. 3. Specimen A, 16 strands of 1 percent Cd 99 percent Cu alloy No. 12, with Fiberglas core. Maximum outside diameter 0.040 inch. 4. Specimen B, Kingston Prod. Corp. Type RX771. Same as Specimen A but with 7-strand stainless steel core. 5. Specimen C, 0.040 inch (No. 18 A.W.G.) copper clad wire. 6. Theoretical resistance of No. 18 A.W.G. copper wire. 7. 0.153-inch 24S aluminum alloy, solid. Measured values agree with theoretical values.

[19] A. Alford and A. G. Kandoian, "UHF Loop Antennas," Trans. A. I. E. E. 59, 846 (1940).
[20] S. Butterworth, "On the AC Resistance of Solenoidal Coils," Proc. Roy. Soc. 107A, 693 (1925).

Specimen C, 22.4 ohms/1000 feet,
Specimen A, 27.0 ohms/1000 feet,
Specimen D, 0.153 AL 24S, 0.88 ohm/1000 feet.

c. Calculated Radiofrequency Resistances

By means of skin effect formulas, the radiofrequency resistance of No. 18 copper wire (which has the same diameter as Specimen C) was calculated for various frequencies, and plotted in Fig. 8, for comparison with experimental results. Likewise, this was done for Specimen D, 0.153 Al 24S wire, based upon the direct current resistance measurement.

RESULTS AND CONCLUSIONS

The transmission line and Q-meter method of measuring the radiofrequency resistance of wires appears to be more reliable than the use of coils, assuming all precautions just described to be taken, and considering test equipment commercially available at present. Also, the results of this method agree closely with those of calculations by skin effect formulas.

For round wires known to consist of homogeneous material, calculated values of radiofrequency resistance appear to be just as reliable and accurate as experimental values, and are much more easily obtained.

ACKNOWLEDGMENT

The author was assisted considerably in the theoretical work by Miss Carolyn B. Parker, and in the mechanical work by Mr. Sergio Brunicardi.

The Absorption Displacement in X-Ray Diffraction by Cylindrical Samples

B. E. Warren

Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received May 25, 1945)

In the precision determination of lattice constants from back reflection Debye-Scherrer patterns, one of the systematic errors arises from the effect of absorption in the sample. The magnitude and form of the error is calculated for a cylindrical sample, for both parallel and diverging primary beams, and for a representative variety of values of the absorption coefficient. Calling 2φ the back scattering angle ($\varphi=90-\theta$), it is found that the error in φ caused by absorption can be set proportional to φ. The form of the absorption correction allows it to be lumped in with the other systematic errors for a linear extrapolation according to either $\Delta d/d = K\varphi \tan\varphi$ or $\Delta d/d = K \sin^2\varphi$. There is some preference for the former.

I. INTRODUCTION

PRECISION determinations of lattice constants are readily made from Debye-Scherrer powder patterns, or single crystal rotation patterns, using back reflections (θ near 90°), and extrapolating out the systematic error. For the Debye-Scherrer powder pattern, the form of the systematic errors has been discussed by Bradley and Jay[1] and by Cohen.[2] From the Bragg law $\lambda=2d\sin\theta$, the fractional error in d caused by an error in θ is given by

$$\frac{\Delta d}{d}=-\frac{\cos\theta}{\sin\theta}\Delta\theta=\frac{\sin\varphi}{\cos\varphi}\Delta\varphi, \qquad (1)$$

where $\varphi=\pi/2-\theta$.

[1] A. J. Bradley and A. H. Jay, Proc. Phys. Soc. **44**, 563 (1932).
[2] M. U. Cohen, Rev. Sci. Inst. **6**, 68 (1935).

The principal systematic errors in φ arise from

1. Shrinkage of the film.
2. Uncertainty in the film radius R.
3. Displacement of the sample from the center of the camera.
4. Effect of absorption in the sample.

Following Bradley and Jay, and Cohen the error $\Delta\varphi_{SR}$ owing to shrinkage and radius uncertainty is given by

$$\Delta\varphi_{SR}=(\Delta S/S-\Delta R/R)\varphi, \qquad (2)$$

where S is the measured distance between corresponding lines on a back reflection film, $\Delta S/S$ is the fractional film shrinkage, and $\Delta R/R$ is the fractional error in the radius. The error $\Delta\varphi_D$ owing to displacement of the sample from the center of the camera is given by

$$\Delta\varphi_D=(\Delta X/R)\sin\varphi\cos\varphi, \qquad (3)$$

where ΔX is the component of displacement parallel to the primary beam. No error is produced by a small lateral component of displacement. The error $\Delta\varphi_A$ caused by absorption in the sample has been given by Bradley and Jay as

$$\Delta\varphi_A = -\left(\frac{1}{R}+\frac{1}{AX}\right)\frac{r\sin 2\varphi}{2(\pi-2\varphi)}. \qquad (4)$$

This is an approximate equation developed for a primary beam which diverges from a point at a distance AX from the sample. The sample is a cylinder of radius r, and absorption in the sample is so high that diffraction takes place only in the surface. The intensity of diffracted radiation is not symmetrical across the beam, and it is assumed that measurement is made to the center of gravity of the diffracted intensity.

Combining the above equations we obtain

$$\frac{\Delta d}{d}=\frac{\sin\varphi}{\cos\varphi}\left\{\left(\frac{\Delta S}{S}-\frac{\Delta R}{R}\right)\varphi + \left(\frac{\Delta X}{R}-\left[\frac{1}{R}+\frac{1}{AX}\right]\frac{r}{\pi-2\varphi}\right)\sin\varphi\cos\varphi\right\}.$$

Since only back reflections are used, where φ is small, φ can be replaced by $\sin\varphi\cos\varphi$ and we have

$$\Delta d/d = K\sin^2\varphi = K\cos^2\theta. \qquad (5)$$

The fractional error in a lattice constant due to systematic errors is accordingly proportional to $\sin^2\varphi$. If the values of a lattice constant, determined from several back reflection lines, are plotted against $\sin^2\varphi$, the points should fall on a straight line. By extrapolating to $\sin\varphi=0$, we obtain the correct value free of all systematic error. This is a simple and beautiful method for precision determination of lattice constants. It is obvious that the possibility of using this method depends upon the fact that all of the various systematic errors have approximately the same dependence upon φ.

Of the various systematic errors, the only one which is not simple and obvious is the error caused by absorption. The absorption error used above is attributed to Bradley and Jay, and is an approximate result for the limiting case of very high absorption. The angular dependence is of the form

$$\Delta\varphi_A \sim \frac{\sin\varphi\cos\varphi}{(\pi-2\varphi)} = \frac{\sin\theta\cos\theta}{2\theta}.$$

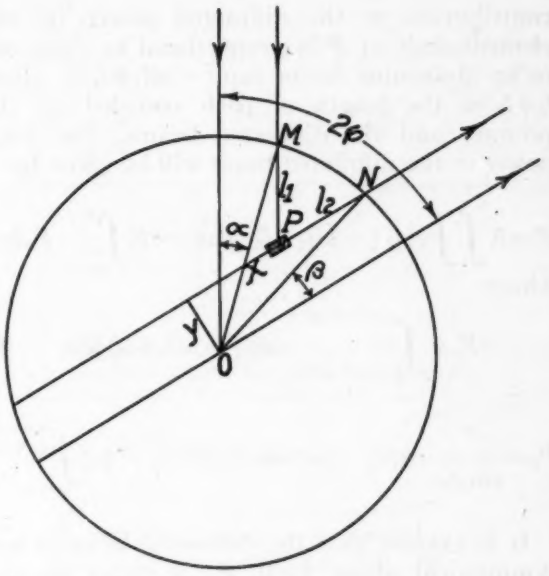


FIG. 1. Path lengths and diffracting volume in a cylindrical sample with parallel primary beam.

The correctness of this form of the absorption error has been questioned recently by Buerger[3] who has derived an absorption correction of the form

$$\Delta\varphi_A \sim \sin^2\varphi = \cos^2\theta.$$

This result is disconcerting, since the form of the absorption correction is not similar to the other corrections. If the proposed form were correct, it would not be possible to lump all of the systematic errors together and extrapolate them out by a linear extrapolation.

It is the purpose of this paper to investigate rigorously the form of the absorption correction, and to show that it has an angular dependence sufficiently like the approximate expression of Bradley and Jay, to allow it to be lumped in with the other corrections for a linear extrapolation. In Sections II–V we shall consider the absorption effect for a parallel primary beam. In Section VI we shall consider the effect for a primary beam which diverges from a point.

II. ABSORPTION DISPLACEMENT FOR A PARALLEL PRIMARY BEAM

We assume that the sample is a perfect cylinder of radius r, and effective linear absorption coefficient μ. The sample can be either a powder or a single crystal. The back scattering angle is 2φ where $2\varphi=\pi-2\theta$. Referring to Fig. 1, the

[3] M. J. Buerger, *X-Ray Crystallography* (John Wiley and Sons, Inc., New York, 1942), Chapter XX.

contribution to the diffracted power by the element $dxdy$ at P is proportional to $dxdy$ and to an absorption factor $\exp[-u(l_1+l_2)]$, where l_1+l_2 is the length of path traveled by the primary and the diffracted beams. The total power in the diffracted beam will be given by

$$P=K\int\int \exp[-\mu(l_1+l_2)]dxdy=K\int_{y=-r}^{y=+r} P_y dy,$$

where

$$P_y=\int_{x=-(r^2-y^2)^{\frac{1}{2}}}^{x=+(r^2-y^2)^{\frac{1}{2}}} \exp[-\mu(l_1+l_2)]dx. \quad (6)$$

$P_y dy$ is proportional to the diffracted power between distances y and $y+dy$ from what would be the center of the diffracted beam if there were no absorption. In terms of the angles α and β shown in Fig. 1, we have the following relations:

$$l_1=r\cos\alpha-y\sin 2\varphi-x\cos 2\varphi,$$
$$l_2=r\cos\beta-x,$$
$$y=r\sin\beta,$$
$$x\sin 2\varphi-y\cos 2\varphi=r\sin\alpha,$$
$$dx=(r/\sin 2\varphi)\cos\alpha d\alpha.$$

Equation (6) becomes

$$P_y=\frac{r}{\sin 2\varphi}\exp[-(\mu r/\sin\varphi)\sin(\varphi-\beta)]\int_{\sin\alpha=-\sin(2\varphi+\beta)}^{\sin\alpha=\sin(2\varphi-\beta)}\exp[-(\mu r/\sin\varphi)\sin(\varphi-\alpha)]\cos\alpha d\alpha. \quad (7)$$

It is evident that the diffracted beam is not symmetrical about $y=0$, P_y is larger on the small φ side (y positive) as indicated in the curves of Fig. 2 and Fig. 4. To what point in the diffracted beam would one measure? The curve P_y plotted against y does not represent the observable intensity distribution across the beam since each element of this curve is spread out by the various sources of broadening which have been neglected. *Assuming a symmetrical law of broadening in which each element of the P_y curve is spread out equally in both directions, the center of gravity of the observable curve will be at the* c.g. *of the P_y curve.* In practice the observable curve usually seems to be broadened enough to make it appear nearly symmetrical. Hence in measuring to the peak of the observable curve, one is measuring very closely to the c.g. of the observable curve, and this is the c.g. of the P_y curve. *We will assume that all measurements are made to the* c.g. *of the observable curve and that this corresponds to the* c.g. *of the P_y curve.* The c.g. of the P_y curve is displaced from the mid-point of the curve for no absorption, by an amount $\bar{y}$ given by

$$\bar{y}=\frac{\int_{-r}^{+r}P_y y dy}{\int_{-r}^{+r}P_y dy}=\frac{\int_0^r (P_y-P_{-y})y dy}{\int_{-r}^{+r}P_y dy}. \quad (8)$$

III. APPROXIMATION FOR VERY SMALL ABSORPTION AND PARALLEL BEAM

For very small values of μr, the exponentials in Eq. (7) can be expanded, giving

$$P_y=\frac{r}{\sin 2\varphi}\int_{\sin\alpha=-\sin(2\varphi+\beta)}^{\sin\alpha=\sin(2\varphi-\beta)}\left\{1-\frac{\mu r}{\sin\varphi}\sin(\varphi-\beta)-\frac{\mu r}{\sin\varphi}\sin(\varphi-\alpha)\right\}\cos\alpha d\alpha. \quad (9)$$

Evaluating the integral and taking the difference P_y-P_{-y}, all terms involving the ambiguities in the limits drop out and we obtain

$$P_y-P_{-y}=2ur^2\sin 2\varphi\sin 2\beta. \quad (10)$$

Combining (8) and (10) and letting $y=r\sin\beta$

$$\bar{y}=\frac{\int_0^{\pi/2}\mu r^4\sin 2\varphi\sin^2 2\beta d\beta}{\int_{-r}^{+r}P_y dy}=\frac{\mu r^2\sin 2\varphi}{4A_{90}}, \quad (11)$$

where A_{90} is the tabulated[4] absorption factor for

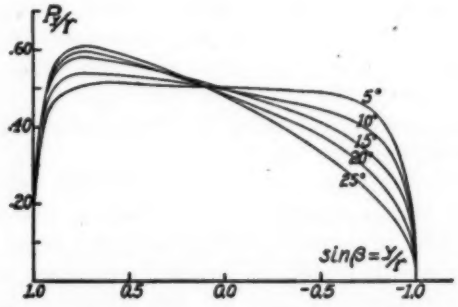


FIG. 2. Values of P_y/r for $\mu r=1.00$ obtained by numerical integration of Eq. (7).

[4] *International Tables for the Determination of Crystal Structures* (Gebrüder Borntraeger, Berlin, 1935) or (Edwards Brothers, Inc., Ann Arbor, Michigan, 1944), p. 584.

$\theta=90°$, for small μr and small φ the factor A is practically independent of φ. For $\mu r=0.2$ (corresponding to a graphite sample 0.4-mm diameter with Cu $K\alpha$-radiation) $\bar{y}/r=0.069 \sin 2\varphi$. The curve for $\mu r=0.2$ is shown in Fig. 6.

IV. INTERMEDIATE ABSORPTION AND PARALLEL BEAM

For intermediate values of μr, Eq. (7) is evaluated by numerical integration. The integration has been carried out for $\mu r=1.00$; $\varphi=5°$, 10°, 15°, 20°, 25°; and about ten different values of β for each value of φ. The results are shown in Fig. 2 where P_y/r is plotted against y/r. The position of the c.g. of each of these curves is listed in Table I.

The values of Table I are shown in Fig. 6 by the curve labelled $\mu r=1.00$.

The dependence of $\bar{y}/r$ upon φ can also be obtained for intermediate values of μr from the results of Rusterholz.[5] In Tables 3 and 4 of his

TABLE I. Center of gravity of the P_y curves of Fig. 2.

φ	$\bar{y}/r$
5°	0.033
10°	0.068
15°	0.106
20°	0.142
25°	0.176

paper, Rusterholz gives the values of $(\mu r)(P_y/r)$ for $\mu r=5.24$ and 8.56 obtained by a numerical integration. The values of (P_y/r) for $\mu r=5.24$ have been plotted against y/r for $\varphi=9.6°$, 21.6°, and 31.9°. From the center of gravity of the areas under these curves the values of $\bar{y}/r$ are obtained. The results are shown in Fig. 6 by the curve labeled $\mu r=5.24$.

V. APPROXIMATION FOR VERY LARGE ABSORPTION AND PARALLEL BEAM

For very high absorption there is no contribution to the diffracted power except from points P (Fig. 1) close to the surface. The distances l_1 and l_2 become so small that the figure MNP can be replaced by a triangle MNP with the side MN tangent to the circle at N. In the terminology of Fig. 1

$$l_1=l_2(\cos\beta)/(\cos(2\varphi-\beta)),$$

$$l_1+l_2=2l_2[\cos\varphi\cos(\varphi-\beta)]/[\cos(2\varphi-\beta)].$$

[5] A. Rusterholz, Helv. Phys. Acta 4, 68 (1931).

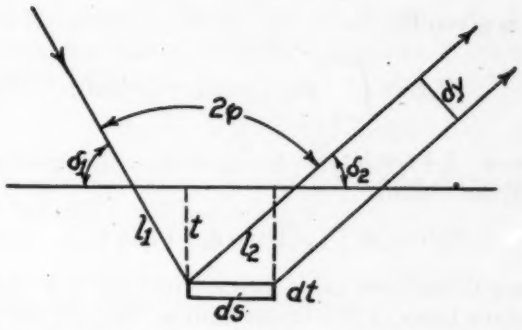


FIG. 3. Path lengths and diffracting volume for diffraction by a planar sample with very high absorption.

Equation (6) becomes

$$P_y=\int_0^\infty \exp\left[-2\mu l_2\frac{\cos\varphi\cos(\varphi-\beta)}{\cos(2\varphi-\beta)}\right]dl_2$$

$$=\frac{\cos(2\varphi-\beta)}{2\mu\cos\varphi\cos(\varphi-\beta)}. \quad (12)$$

Since Eq. (12) is fundamental to the rest of this paper, it is worth while to consider its derivation from another point of view. For high absorption, diffraction takes place only in the surface layer, and a small area on the cylindrical surface can be approximated by a planar sample. In the notation of Fig. 3, the diffracted power from the region under the surface area of length

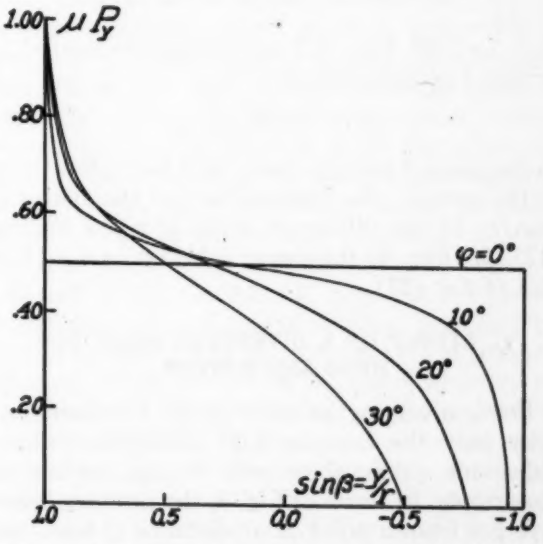


FIG. 4. Distribution of energy in the diffracted beam for very high absorption. Plot of Eq. (12).

ds is given by

$$P_s ds = \int_{t=0}^{t=\infty} \exp\left[-\mu(l_1+l_2)\right] ds\, dt, \qquad (13)$$

where $l_1+l_2=t(1/\sin\delta_1+1/\sin\delta_2)$. Integrating (13) we obtain

$$P_s ds = [ds]/[\mu(1/\sin\delta_1+1/\sin\delta_2)]. \qquad (14)$$

Since diffraction takes place within a very thin surface layer of thickness t, all of this diffracted power is confined within a beam width dy, where

$$\frac{dy}{ds}=\sin\delta_2, \quad P_s ds = P_y dy.$$

Making these substitutions in (14) we obtain

$$P_y = \frac{\sin\delta_1}{\mu(\sin\delta_1+\sin\delta_2)}. \qquad (15)$$

In the notation of Fig. 1, $\delta_1=90°-(2\varphi-\beta)$ and $\delta_2=90°-\beta$. Inserting these values into (15) we obtain Eq. (12).

In Fig. 4 the quantity μP_y as given by Eq. (12) is plotted against $y/r=\sin\beta$ for $\varphi=0°$, 10°, 20°, and 30°. The marked assymmetry in the reflected beam and the cut-off on the large φ-side at $\beta=-(90-2\varphi)$ are clearly seen. The center of gravity of the diffracted power can be found analytically from Eq. (12).

$$\bar{y}=\frac{\int P_y y\,dy}{\int P_y dy}=\frac{r\int_{-(\pi/2-2\varphi)}^{+\pi/2}\frac{\cos(2\varphi-\beta)}{\cos(\varphi-\beta)}\sin\beta\cos\beta\,d\beta}{\int_{-(\pi/2-2\varphi)}^{+\pi/2}\frac{\cos(2\varphi-\beta)}{\cos(\varphi-\beta)}\cos\beta\,d\beta}. \qquad (16)$$

Let $\varphi-\beta=\epsilon$

$$\bar{y}/r=\frac{\int_{\pi/2-\varphi}^{\varphi-\pi/2}(\cos 2\varphi+\cos 2\epsilon)(\sin\varphi-\cos\varphi\tan\epsilon)\,d\epsilon}{\int_{\pi/2-\varphi}^{\varphi-\pi/2}2\left(\frac{\cos^2\epsilon-\sin^2\varphi}{\cos\epsilon}\right)d\epsilon}.$$

From which we obtain

$$\bar{y}/r=\frac{\sin\varphi}{4}\left\{\frac{(\pi-2\varphi)\cos 2\varphi+\sin^2 2\varphi}{\cos\varphi+\sin^2\varphi\ln\tan(\varphi/2)}\right\}. \qquad (17)$$

For $2\varphi<50°$ Eq. (17) can be approximated to within 2 percent accuracy by

$$\bar{y}/r=(\pi/4)\tan\varphi. \qquad (18)$$

For a parallel primary beam and high absorption in the sample, the displacement of the center of gravity of the diffracted beam is given by Eq. (17). In Fig. 6, the curve labelled $\mu r=\infty$ is a plot of Eq. (17).

VI. EFFECT OF A DIVERGING BEAM FOR HIGH ABSORPTION

For a diverging primary beam we shall consider only the case for high absorption where diffraction takes place only in the surface of the sample. In terms of Fig. 5, the primary beam diverges from a point X at distance D from the sample, r is the radius of the cylindrical sample, R is the film radius, and 2φ is the back scattering angle. That part of the primary beam striking the sample at P, makes an angle γ with the central line XO, where

$$\gamma=[r\sin(2\varphi-\beta)]/D. \qquad (19)$$

The primary and diffracted beams at P make angles δ_1 and δ_2 with the tangent at P

$$\delta_1=90-(2\varphi-\beta+\gamma), \quad \delta_2=90-(\beta-\gamma).$$

Putting these values in (15) we obtain

$$P_y=\frac{\cos(2\varphi-\beta+\gamma)}{2\mu\cos\varphi\cos(\varphi-\beta+\gamma)}.$$

Since in practice γ is a very small angle, it is a sufficiently good approximation to write

$$P_y=\frac{\cos(2\varphi=\beta)}{2\mu\cos\varphi\cos(\varphi-\beta)}. \qquad (20)$$

Let O' be the point where diffracted radiation from the center of the sample would strike the film. The diffracted radiation from P strikes the

film at a distance S from O' where

$$S=r\sin\beta+R\gamma=r\left\{\sin\beta+\frac{R}{D}\sin(2\varphi-\beta)\right\}. \quad (21)$$

Let $P_S dS$ be the diffracted power striking the film between distances S and $S+dS$ due to a particular region P on the sample. (At a particular point S on the film there may be contributions from two different points P on the sample.) Then

$$P_S dS=P_y dy=P_y\cos\beta d\beta. \quad (22)$$

The position on the film of the center of gravity of the diffracted beam is given by

$$\bar{S}=\frac{\int P_S S dS}{\int P_S dS}=\frac{r\int_{-(\pi/2-2\varphi)}^{+\pi/2}\frac{\cos(2\varphi-\beta)}{\cos(\varphi-\beta)}\left\{\sin\beta+\frac{R}{D}\sin(2\varphi-\beta)\right\}\cos\beta d\beta}{\int_{-(\pi/2-2\varphi)}^{+\pi/2}\frac{\cos(2\varphi-\beta)}{\cos(\varphi-\beta)}\cos\beta d\beta}. \quad (23)$$

In the above equation the approximation that γ is small has been involved in both the value of P_y and in the limits of integration. By making the same change in variable as used in evaluating (16), namely, $\varphi-\beta=\epsilon$ the integration in (23) can be carried out and we obtain

$$\bar{S}=\left(1+\frac{R}{D}\right)\frac{r\sin\varphi}{4}\times\left\{\frac{(\pi-2\varphi)\cos 2\varphi+\sin 2\varphi}{\cos\varphi+\sin^2\varphi\ln\tan(\varphi/2)}\right\}. \quad (24)$$

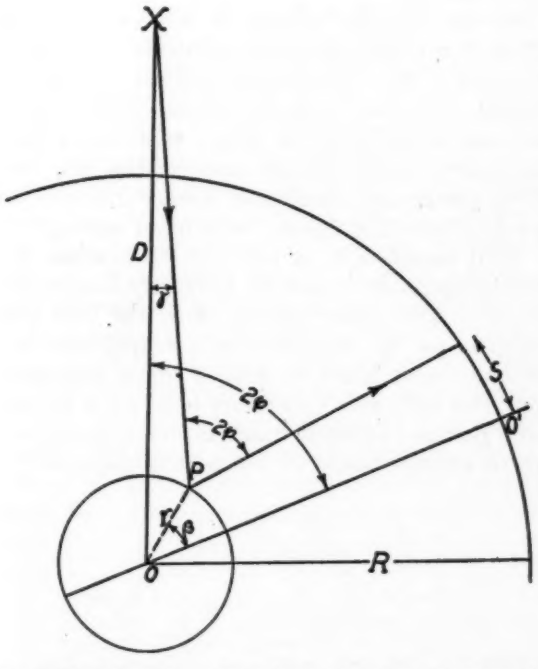


FIG. 5. Path lengths and angles involved in the discussion of absorption with a diverging primary beam.

Comparing this result with (17) it is seen that for high absorption, the angular dependence is the same for either the parallel or the diverging primary beam.

$$\bar{S}=(1+R/D)\bar{y}. \quad (25)$$

VII. CONCLUSIONS

The error in φ caused by absorption, which we have called $\Delta\varphi_A$, is given by

$$\Delta\varphi_A=\frac{\bar{y}}{2R}\text{ parallel primary beam,}$$
$$\Delta\varphi_A=\frac{\bar{S}}{2R}\text{ diverging primary beam.} \quad (26)$$

For a parallel primary beam the curves of Fig. 6 give the relation between $\bar{y}$ and φ. It is evident that $\bar{y}$ is approximately proportional to φ, and hence from (26) $\Delta\varphi_A$ is approximately proportional to φ. The functional dependence varies slightly with the magnitude of μr. For small μr, $\Delta\varphi_A$ is proportional to $\sin 2\varphi$, for intermediate values of μr it is proportional to φ, and for large values of μr it is proportional to $\tan\varphi$. In practice the value of μr will usually fall between 1.00 and ∞. Hence there is some preference for setting $\Delta\varphi_A$ proportional to φ rather than $\sin 2\varphi$.

It is evident from the curves of Fig. 6 that for most of the values of μr which we meet in practice, the absorption effect is practically equivalent to that of the limiting case of very high absorption. In Section VI for a diverging beam we treated only the case of very high absorption. From the above discussion this will suffice to cover the majority of cases for a diverging primary beam. For high absorption

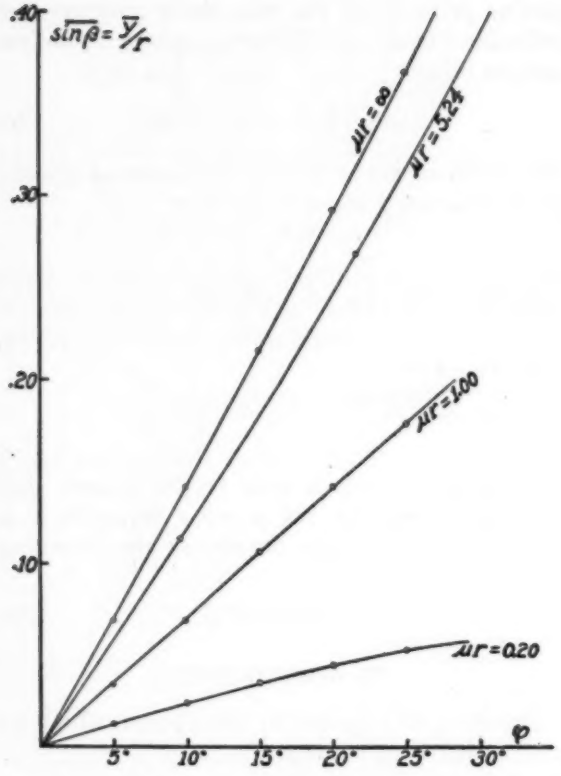


FIG. 6. Displacement of the center of gravity of the diffracted beam by absorption for a parallel primary beam. $\mu r = 0.20$ is a plot of Eq. (11), $\mu r = 1.00$ is a plot of the values of Table I obtained by numerical integration, $\mu r = 5.24$ is obtained from the results of Rusterholz, and $\mu r = \infty$ is a plot of Eq. (17).

we found the same angular dependence for either a diverging or a parallel primary beam. *Hence for small values of φ and either a diverging or a parallel primary beam we can set $\Delta\varphi_A$ proportional to φ.*

The errors in φ caused by film shrinkage, radius uncertainty, and absorption are all proportional to φ. The error caused by sample displacement is proportional to $\sin\varphi\cos\varphi$, but for small φ this can be approximated as φ. Hence the systematic errors can all be lumped together in the form $\Delta\varphi = K\varphi$ and Eq. (1) becomes

$$\Delta d/d = K\varphi \tan\varphi. \tag{27}$$

While there is a slight preference for the form of Eq. (27), it is obvious that for small φ an essentially equivalent form is

$$\Delta d/d = K \sin^2\varphi. \tag{28}$$

This is the answer to the question posed at the beginning of the paper. *The form of the absorption correction is sufficiently like the approximate expression of Bradley and Jay, to allow it to be lumped in with the other corrections for the Bradley and Jay, and Cohen type of linear extrapolation.*

For the usual case of high absorption, a comparison of Eqs. (3), (18), (25), and (26) shows that for small φ the effect of absorption is equivalent to a displacement of the sample toward the entrance slit by amounts

$$\Delta X = \frac{\pi}{8} r \qquad \text{Parallel primary beam;}$$

$$\Delta X = \frac{\pi}{8} r\left(1 + \frac{R}{D}\right) \qquad \text{Diverging primary beam.}$$

Since in practice the value of μr is usually so large that the absorption effect is nearly equivalent to the limiting effect for very high absorption, it is unnecessary to worry about the contribution by extinction to the effective absorption coefficient.

For the average sample μr is very large, and diffraction takes place essentially in the surface. In general the sample will not be the perfect cylinder with the perfectly smooth surface which has been assumed in the above treatment. Small departures could change appreciably the form of the absorption correction. Hence it is not safe to rely absolutely upon a treatment such as this of ideal samples. It is not safe to assume that the extrapolation must be perfectly linear. The results of this paper merely show the best variables to use in making the extrapolation and the ones most likely to give a linear extrapolation. The only safe procedure is to have at least three points and then make either a linear or a curved extrapolation as the points indicate.

Molecular Weight Studies on High Polymers with the Electron Microscope

R. F. Boyer and R. D. Heidenreich
The Dow Chemical Company, Midland, Michigan

(Received June 15, 1945)

Some preliminary results are reported on measurements of molecular weight distributions for polychlorostyrene. The following technique of isolating single polymer molecules is used: A precipitating agent, such as propanol, is added to a very dilute (1 ppm) solution of the polymer in benzene, presumably causing the molecules to coil up. Drops of this mixture are evaporated on a silica screen and examined in the microscope. Circular particles ranging in diameter from 15A to 500A and believed to represent single polymer molecules have been obtained on various polymers. The relationship between measured particle diameter and molecular weight is reviewed in some detail. The data appear to favor the concept of a random coil to the extent that particle diameter varies as the square root of the molecular weight. The molecular weight distribution curves obtained are of the expected shape and extent, although the average molecular weights computed from the distribution curves are 4–5 times greater than values measured by an independent method. Finally, the role of electrical charges in stabilizing isolated polymer molecules is discussed.

INTRODUCTION

CURRENT scientific and technological interest in average molecular weights and molecular weight distributions of linear or slightly branched (soluble) high polymers has prompted an investigation of this subject using the electron microscope technique. Such an approach is predicated on the hope of finding simple methods of isolating single polymer molecules in a form suitable for microscope examination. This report will outline some of the problems involved and present some preliminary results. For example, exploratory questions such as attainable resolving power of the electron microscope for organic molecules; types of polymers, solvents, and precipitating agents; correct concentration ranges; shape and size of chain molecules; and general techniques are reviewed.

It is recognized that the general problem of absolute molecular weight determinations on high polymers is now in a rather favorable state. Refinements in technique of osmotic pressure measurements recently introduced by Fuoss,[1] Flory,[2] Julander and Svedberg,[3] as well as enhanced theoretical treatments of the osmotic pressure problem by Flory,[4] Huggins,[5] Miller,[6] and Orr[7] have contributed to this situation. Moreover, the light scattering method appears to furnish another reliable absolute determination of molecular weight.[8,9] In fact, it is especially sensitive for large molecules, it is well adapted to low concentration ranges where high polymers exhibit more ideal behavior, and it gives an indication of the breadth of the molecular weight distribution curve.[10] This naturally raises doubts as to the value of further methods.

An electron microscope method, if successful, should give the entire molecular weight distribution curve, from which various average molecular weights can be computed. Moreover, present indications are that it only begins to show results at concentrations some three orders of magnitude lower than the lowest osmotic pressure concentrations, and several magnitudes lower than can be used in the light scattering method. These two possibilities, then, of exploring an entirely new region of polymer concentrations, and of obtaining complete distributions on heterogeneous polymers are sufficiently attractive to justify some study. Finally, there is the possibility of learning something about the shapes of polymer molecules in solution.

The first systematic work on molecular weights of polymers with the electron microscope is apparently that of Husemann and Ruska.[11] They deposited aqueous solutions of glycogen, in concentrations ranging from 1 to 0.001 mg per cc, onto collodion films which were then examined

[1] R. M. Fuoss and D. J. Mead, J. Phys. Chem. **47**, 59 (1943).
[2] P. J. Flory, J. Am. Chem. Soc. **65**, 372 (1943).
[3] I. Julander and T. Svedberg, Nature **153**, 523 (1944).
[4] P. J. Flory, J. Chem. Phys. **10**, 51 (1942).
[5] M. L. Huggins, J. Am. Chem. Soc. **64**, 1712 (1942).
[6] A. R. Miller, Proc. Camb. Phil. Soc. **39**, 54 (1943).
[7] W. J. C. Orr, Trans. Faraday Soc. **40**, 320 (1944).
[8] P. Debye, J. App. Phys. **15**, 338 (1944).
[9] P. M. Doty, B. H. Zimm, and H. Mark, J. Chem. Phys. **12**, 144 (1944); **13**, 159 (1945).
[10] B. H. Zimm and P. M. Doty, see reference 9, p. 203.
[11] E. Husemann and H. Ruska, J. f. prakt. Chemie **156**, 1 (1940).

after the solutions had dried. A glycogen fraction having an osmotic molecular weight of 1,500,000 gave, especially at the lower concentrations, some roughly circular but not very sharp images ranging from 50 to 150A in diameter. Molecular weights calculated on the assumption that these particles had the same density as the bulk polymer were in good agreement with the osmotic molecular weight. Another fraction of molecular weight 65,000 gave particles that were scarcely discernible. They also[12] prepared an iodine derivative of glycogen containing 40.6 percent iodine in order to obtain more contrast in the images. This sample with a molecular weight of 6,000,000 showed black particles varying from 150 to 300A in size compared with a calculated value of 240A.

More recently Kropa, Burton, and Barnes[13] have investigated films of polystyrene cast onto a water surface from solution in organic solvents. These films exhibited at various points black nodules whose size was in agreement with that predicted by the viscosity molecular weight of the polymers, if one assumed that the material in these nodules had the same bulk density as polystyrene. However, the method did not appear to work for all specimens, but only on certain high molecular weight polymers prepared in a special manner.

SHAPES OF POLYMER MOLECULES IN SOLUTION

The glycogen molecules examined by Ruska and Husemann were very likely spherical in shape as evidenced by the fact that their dilute solution viscosities were only slightly greater than that predicted by the Einstein equation for suspensions of rigid spheres. Solvation or swelling of these particles could easily account for the slightly enhanced viscosity. These molecules were, therefore, of convenient shape for examination with the electron microscope. However, the abnormally high viscosity of dilute solutions of linear polymers (polystyrene, methyl methacrylate, polyvinyl chloride, etc.) suggests that their molecules possess an elongated shape in solution.[14] They are, therefore, not in a favorable configuration for examination by the electron microscope, especially when an attempt is being made to compute their volumes.

In general, linear high polymeric molecules might be characterized by having one dimension (the extended length or the chain contour length) which may reach a micron or more. They would thus be accessible to the light microscope except for the fact that the cross dimension of the chain is only 10–20A units, and thus even below the limit of the electron microscope. Actually, these chain molecules seldom exist in their fully extended state but occur as loose, randomly kinked chains whose extent in space depends on such factors as the solvating power of the solvent, the degree of free rotation about the carbon-carbon bonds, and the space requirements of substituent groups. A single polystyrene molecule of molecular weight 80,000 would possess

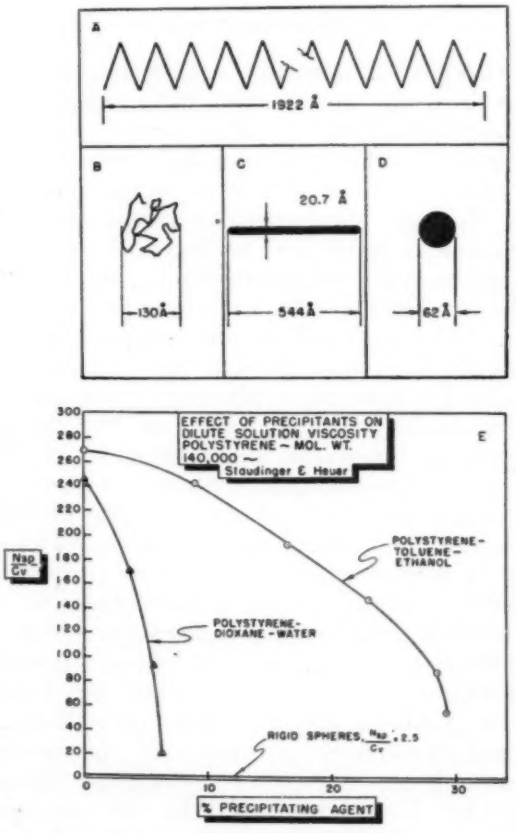


FIG. 1. Shapes and dimensions of a polystrene molecule of molecular weight 80,000. *A* is fully extended, *D* is completely compressed to bulk density, *B* is for a random coil, and *C* is calculated shape from experimental data. Curve *E* shows the decrease in specific viscosity, η_{sp}, divided by original volume concentration of polymer, C_v, on adding a non-solvent. These two curves end abruptly at the precipitation point, which should correspond to configurations such as *B* or *D*.

[12] E. Husemann and H. Ruska, Naturwiss. **28**, 534 (1940).
[13] E. L. Kropa, C. J. Burton, and R. B. Barnes, Electron Microscope Society Meeting, New York City, January 11–15, 1944.
[14] R. Simha, J. App. Phys. **13**, 147 (1942).

a projected length of 1922A in the fully extended state or a chain contour length of roughly 2400A, but would form a compact sphere of 62A-diameter having the same bulk density as polystyrene. Figure 1 illustrates several typical configurations and the appropriate dimensions for a polystyrene molecule of molecular weight 80,000. The dimensions of the highly elongated shape (1-C) were taken from a paper by Simha[14] who based his calculations on ultracentrifuge and viscosity measurements. The diameter of the randomly kinked chain was calculated by a method which will be discussed in detail later. (Equation (9) with $\gamma=4$.)

In solution a linear polymer molecule will assume a shape intermediate between that of the fully uncoiled and highly improbable state (1-A) and the compact sphere of 1-D. Studies of polystyrene solutions by Staudinger and Heuer,[15] and more recently by Alfrey, Bartovics, and Mark[16] have indicated that the viscosity, and presumably the shape, of the molecule can be controlled by proper choice of solvent. Good solvents such as benzene or toluene produce a high viscosity and likely a configuration between 1-A and 1-C, but closer to the latter. Poor solvents such as methyl ethyl ketone, or good solvents to which a precipitating agent like ethyl alcohol has been added, result in low viscosities, and a more compact configuration. In fact, Staudinger and Heuer showed that the addition of water to a dioxane solution of polystyrene eventually reduced the viscosity to a value only a few times greater than the Einstein value for rigid spheres. Such molecules would presumably have shapes intermediate between 1-B and 1-D.

Viscosity studies of this type, which have also been discussed and interpreted in detail by Huggins,[17] furnished the key to our attack on the problem of measuring molecular weights of polymers with the electron microscope. It was proposed that by working at extremely low dilutions either in poor solvents, or in solvent-precipitant systems, the individual chain molecules could be brought into a reasonably compact configuration whose dimensions would be accessible to the electron microscope and whose spherical symmetry would facilitate the calculation of molecular weights from measured particle diameters.

It is convenient to divide this problem into four parts as follows:

1. The ability of the electron microscope to image small particles of organic material.
2. Experimental attempts to isolate single polymer molecules.
3. The relationship between molecular weight and measured particle size.
4. Miscellaneous experiments.

RESOLVING POWER OF ELECTRON MICROSCOPE FOR ORGANIC MOLECULES

The imaging of small particles such as large, single molecules has been described by Von Ardenne.[18–20] He has exhibited micrographs of haemocyanin and glycogen molecules and computed the molecular weights from their dimensions. Some discussion of the limiting resolution obtainable with such materials is presented also, and he concluded that the smallest detectable organic particles are about 40A in diameter when supported on a collodion substrate.

The two main factors encountered in resolving very small particles are the lens defects (chromatic and spherical aberration) and contrast in the image. It is generally agreed[21–23] that the magnetic electron microscope is capable of a resolution of about 10A and this has been demonstrated for gold sols, etc. Hillier[24] has directed attention to the contrast obtainable in the image and computes that it should be ideally possible to image a single atom of atomic number $Z>25$ with 60-kv electrons. For atoms of atomic numbers such as found in most organic molecules the contrast limits the size of particle that can be discerned so that only groups of atoms or molecules can be detected. Hillier's calculations are based on taking the smallest observable difference in plate transmission between the particle and the background as being about 10 percent. Thus, the smallest detectable particle of a given material that can be observed will depend

[14] R. Simha, J. App. Phys. **13**, 147 (1942).
[15] H. Staudinger and W. Heuer, Zeits. f. physik. Chemie **A171**, 129 (1934).
[16] T. Alfrey, A. Bartovics, and H. Mark, J. Am. Chem. Soc. **64**, 1557 (1942).
[17] M. L. Huggins, J. App. Phys. **14**, 246 (1943).
[18] M. v. Ardenne, Zeits. f. physik. Chemie **A187**, 1 (1940).
[19] M. v. Ardenne and D. Beischer, Zeits. f. physik. Chemie **B45**, 465 (1940).
[20] M. v. Ardenne and D. Beischer, Kautschuk **16**, 55 (1940).
[21] M. v. Ardenne, Zeits. f. Physik **108**, 338 (1938); **111**, 152 (1938); **112**, 744 (1939).
[22] M. v. Ardenne, *Elektronen Übermikroskopie* (Verlagsbuchhandlung, Julius Springer, Berlin, 1940).
[23] B. V. Borries and E. Ruska, Zeits. f. tech. Physik **20**, 225 (1939).
[24] James Hillier, Phys. Rev. **60**, 743 (1941).

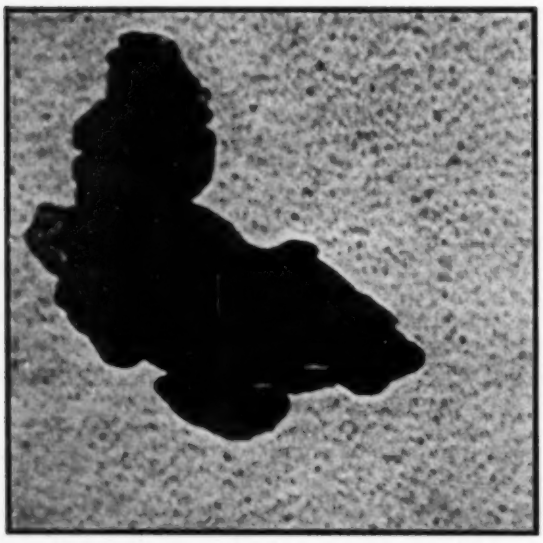


FIG. 2. $Mg(OH)_2$ powder on ethylcellulose substrate. The fine structure in the substrate, believed to represent micellae, illustrates the resolving power obtainable on organic materials. ×100,000.

upon the cross section for scattering presented by that material.

The relation between contrast in the image and the thickness of the object, D, can be obtained from the relation.

$$I = Io \exp(-\beta D), \tag{1}$$

where I is the transmitted intensity, Io, the initial intensity, and β is the absorption coefficient or total cross section for scattering. Marton[25] has employed Eq. (1) to objects thin enough to allow the assumption of single scattering. Values for the elastic and inelastic cross sections for scattering of the commoner atoms can be obtained from his paper and employed in (1) with

$$\beta = \Sigma_i S_i N_i, \tag{2}$$

when S_i is the cross section for scattering of the atoms of type i and N_i is the number of such atoms per cm^3. S_i is the sum of elastic, inelastic, and electron scattering cross sections. The cross section for elastic scattering is generally the dominant term and varies as $Z^{4/3}$ where Z is the atomic number. Thus the contrast in the image is quite sensitive to atomic number as well as to variations in thickness. In the case of crystalline materials[26–28] the image intensity is complicated by interference phenomena, and Eq. (1) is no longer adequate although for amorphous materials it can be employed at least semi-quantitatively depending upon the data on scattering cross sections that is available.

The contrast in the image is also affected by the condenser and objective apertures according to Borries and Ruska.[29] Increasing the condenser aperture reduces the contrast for constant objective aperture as does increasing the objective aperture for constant condenser aperture. All micrographs in this paper were taken with an RCA electron microscope, Type B with a platinum objective limiting aperture. The necessity of this aperture has been questioned, and some instruments are operated with the limiting aperture and others without. It has been the experience of this laboratory that the objective limiting aperture aids appreciably in obtaining contrast. Figure 2 is an example of the image quality that can be obtained. This micrograph shows finely ground $Mg(OH)_2$ powder supported on an ethyl cellulose substrate. Of chief interest is the fine structure exhibited in the ethyl cellulose substrate. This film was cast on water from a lacquer solution containing excess alcohol, yielding a granular film not usually obtained at lower alcohol concentrations. The small spots are probably micellae or regions of crystallinity with a size of 40–100A and are uniformly distributed through the entire film. The smaller particles are at the limit of resolution computed by Von Ardenne.[18]

In this connection it should be mentioned that the photographic technique is important in imaging small particles. Generally satisfactory results are obtained by using contrast lantern slide plates and D-8 developer. The negatives are enlarged and printed on contrast paper with particular care as to density and contrast. The electron exposures are made with the viewing screen intensity reduced such that the brightest part of the field is just visible requiring an exposure time of about 13 seconds. Plates so exposed and processed can be enlarged 10–20 times without difficulties with graininess. Blanks have been run and the validity of the structures shown in the remainder of the illustrations ascertained.

When the particles or molecules to be studied are not to be part of the actual film but rather are to be isolated units, they must be supported by a substrate of some kind. The supporting

[25] L. Marton, J. App. Phys. **12**, 759 (1941).
[26] R. D. Heidenreich, Phys. Rev. **62**, 291 (1942).
[27] B. V. Borries and E. Ruska, Naturwiss. **28**, 366 (1940).
[28] R. D. Heidenreich and L. Sturkey, J. App. Phys. **16**, 97 (1945).
[29] B. V. Borries and E. Ruska, Zeits. f. tech. Physik **19**, 402 (1938).

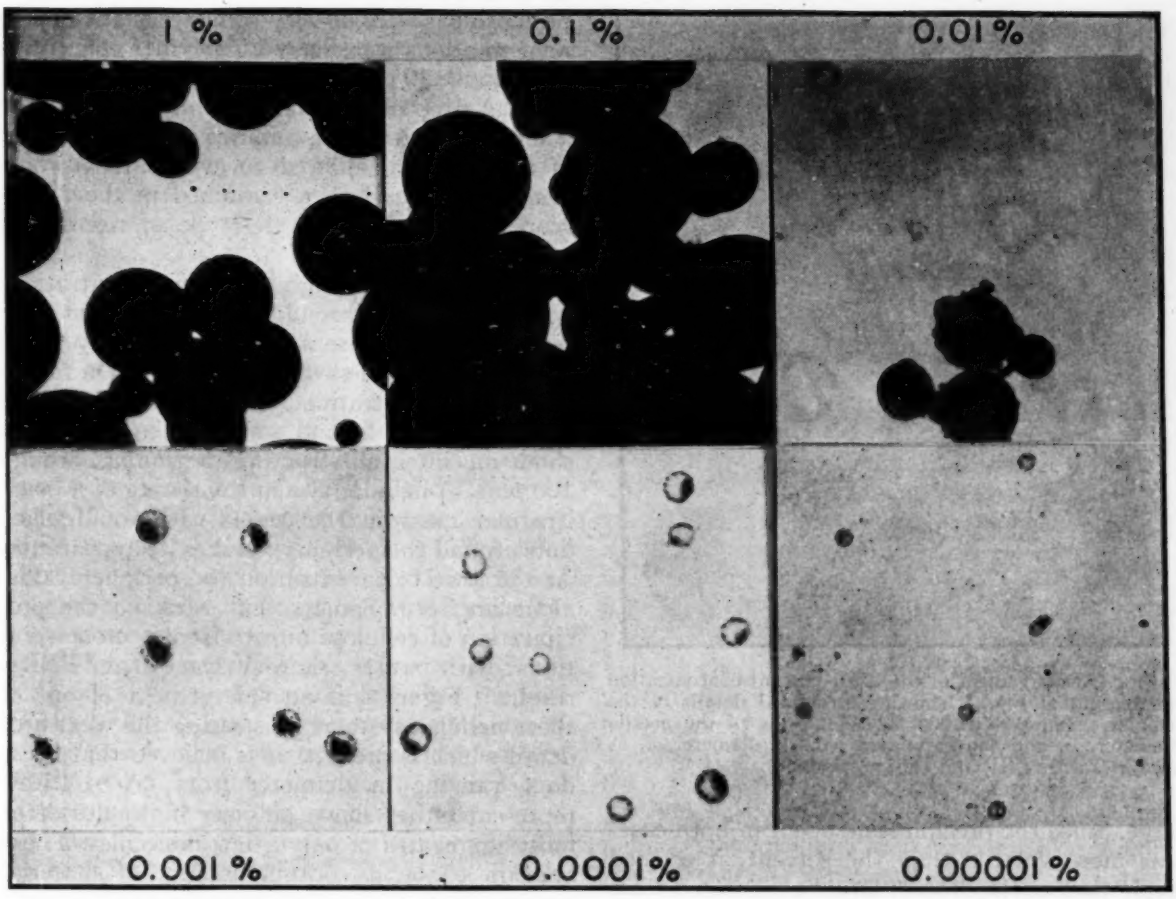


FIG. 3. Structures obtained by adding five volumes of *n*-propanol to benzene solutions of polychlorostyrene of the indicated concentrations in percent. ×16,000.

film obviously is a disadvantage as far as resolution is concerned and should be as transparent as possible. Plastic films such as cellulose nitrate, Formvar, or ethyl cellulose have been commonly used as supports for general microscope work. Since the molecules to be employed in this work are deposited from solvents which readily dissolve the films mentioned, it was necessary to use inorganic films. One such film is an anodic aluminum oxide film.[30,31] They are inclined to exhibit some granular structure of the order of 100–200A and are difficult to prepare completely free of foreign material. It was found that the best support films for this work were the silica films employed in replica work.[32] The silica films were prepared by first coating a clean glass microscope slide with a low viscosity polystyrene lacquer film. Silica was then evaporated onto the slide in a vacuum chamber and removed as in the case of the replicas.

[30] Max A. Lauffer, Chem. Rev. **31**, 561 (1942).
[31] G. Hass and H. Kehler, Kolloid Zeits. **95**, 26 (1941).
[32] R. D. Heidenreich and V. G. Peck, J. App. Phys. **14**, 23 (1943).

EXPERIMENTAL TECHNIQUES

Our first attempt to obtain molecular weights of polymers using the electron microscope technique was by precipitation of a 0.001 percent benzene solution of polystyrene with methanol. When drops of this solution were deposited on silica slides and examined, diffuse spots were obtained having an average diameter of 250A. The general diffuseness of the images was ascribed to the low electron scattering power of a hydrocarbon polymer. A polychlorostyrene sample was, therefore, examined again as a 0.001 percent benzene solution, but this time precipitated with hexane. The images were much blacker but still relatively indistinct and quite

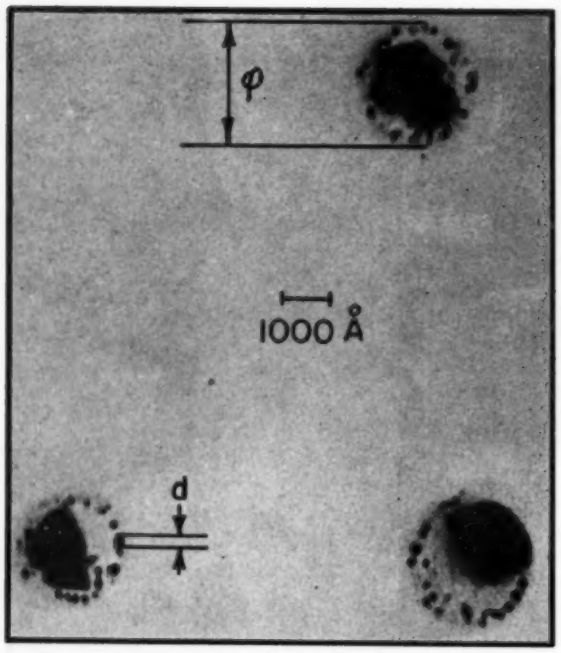


FIG. 4. Enlargement of the 0.001 percent concentration micrograph of Fig. 3, showing structural details of the nebular systems. ×64,000. The definition of the nebula diameter, φ, and the individual particle diameter, d, as used in Tables I, II, and III are given.

large. Since the precipitating agent in both cases was more volatile than the solvent, it seemed likely that as the drop of liquid evaporated from the silica substrate, it would become progressively richer in solvent. This would tend to redissolve the precipitated particles and thus give them a hazy appearance. All future work was, therefore, confined to precipitants which were less volatile than the solvent.

For example, a series of benzene solutions ranging in concentration from 10,000 to 0.1 ppm (1 to 0.00001 percent) were prepared from polyparachlorostyrene having a weight average molecular weight of 140,000 (determined by light scattering). This sample had been polymerized for three days at 125°C to a conversion of 97 percent. The above solutions were then treated in the following manner. One volume of *n*-propyl alcohol was quickly added to one volume of solution, and the mixture heated for two minutes at 80°C. The purpose of this heating was to redisperse any aggregates of molecules which may have formed during the alcohol addition After the mixture had cooled slowly to room temperature, an additional four volumes of *n*-propyl alcohol were added rapidly to extract adsorbed benzene from the polymer and to promote particles of very compact shape. These final solutions were hazy at polymer concentrations above 10 ppm and clear at lower concentrations. A drop of the final solution was placed on a silica film within ten minutes after preparation of the solution, allowed to evaporate at room temperature, and then examined in the microscope. Figure 3 shows the type of results obtained.

It is seen that at high polymer concentrations spherical masses around 1μ in diameter predominate. These are aggregates of many molecules possibly still swollen by solvent. On reducing the concentration, the diameter of these particles decreases and some of them begin to show an internal structure beginning around 100 ppm. A nebular system consisting of a black irregular mass in the center with small black dots around the periphery makes its appearance. At the lowest concentration the peripheral dots alone are seen. Some recent work on the precipitation of cellulose nitrate from acetone solutions with water showed somewhat similar results.[33] Figure 4 is an enlargement of one of these nebular systems illustrating the wealth of detail which is present. It is believed that these dots, ranging in diameter from 60 to 250A, represent either single polymer molecules or at most aggregates of only a few molecules. There appears to be a continuous film of material ranging from one particle to the next. This may represent low molecular weight material which was soluble even in the solvent-precipitant system used, and which gave a more or less coherent film.

It will also be noted that outside of the nebular systems, and scattered over the entire field, are a number of diffuse particles. These again are believed to be molecules of intermediate molecular weight which while not soluble yet may have been highly swollen in the benzene-propanol mixture. It is recognized, of course, that in a heterogeneous polymer containing a wide range of particle sizes, each molecular weight species will have different solubility characteristics. Fractions of this polymer containing much narrower molecular weight bands have been prepared and are under investigation. Ultimately, of course, the ideal situation is to be able to determine the complete molecular weight distribution of a polymer.

[33] G. Hambraens and B. Ranky, Nature 155, 200 (1945). See also D. L. Katz and K. E. Beu, Ind. Eng. Chem. 37, 195 (1945).

We wish now to study in some detail the actual distribution of particle sizes for the material in these nebular systems, and then to attempt some calculations of molecular weight. Before doing this, however, it seems appropriate to speculate briefly concerning the origin of the nebular system. Two hypotheses have been considered thus far with no conclusive evidence for choosing between them.

One of them is based on the idea that during the addition of alcohol to the benzene solution aggregates of molecules are formed at even the lowest concentrations. These aggregates, presumably very rich in benzene, would form weak gels possessing a roughly spherical shape in solution. On aging, solvent would be extracted from the gel, the action naturally occurring first at the boundary between gel and supernatant liquid. It was then thought that single molecules or groups of several molecules would condense out of the gel near its periphery, thus giving rise to the observed structure.

This line of thought suggested some long time aging experiments of the solutions after the final alcohol addition but prior to preparing the microscope slides. Figure 5 represents the very evident aging phenomena which occurred on the 10-ppm solution over a one-week geriod. The first trend was that the nebular structure became cleaner and that the nucleus began to show some details. Later the entire nebular structure was destroyed, and then some large diffuse particles formed by what may have been a von Smoluchowski coagulation process. The 1 percent solution showed absolutely no change in particle size or nature even after a week of aging. This type of behavior appeared to be in accord with the hypothesis mentioned above.

The alternate proposal was that in concentrated solutions aggregation of molecules did occur, but that with increasing dilution the trend was toward isolation of individual, coiled-up molecules. However, while the solution was drying on the silica substrate, there was an exponential increase in concentration during the final stages of drying, thus leading to a rapid von Smoluchowski coagulation. A few large particles, which might be aggregates of molecules, or foreign material, served as a nucleus about which the smaller particles condensed, giving rise to a nebular appearance. Finally, the lowest molecular weight material would be deposited over these particles as a semi-continuous film.

Evidence for this hypothesis is suggested in part by the planar nature of the objects (concentration of particles around the periphery, and occasional strings of particles branching out from the nucleus) which might occur in a two dimensional system. Moreover, somewhat similar nebular systems have been observed from highly diluted polymer emulsions deposited and dried on a silica screen (Fig. 6); and even from the

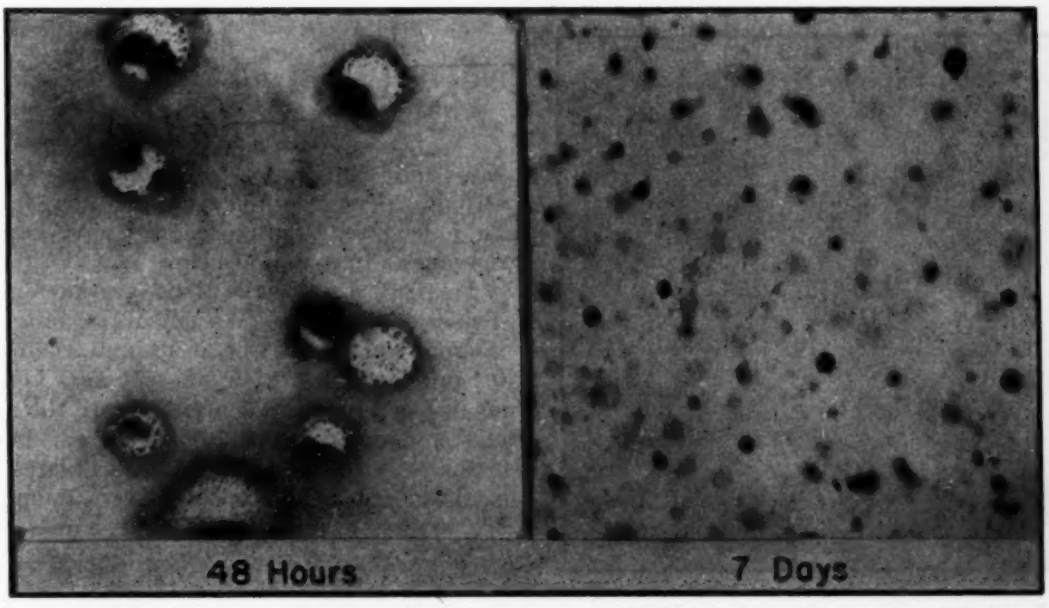


FIG. 5. Aging of the benzene-polychlorostyrene-*n*-propanol solutions prior to casting films. Concentration of polymer was 0.001 percent. ×16,000. The blank with 10 minutes aging time appears in Fig. 3.

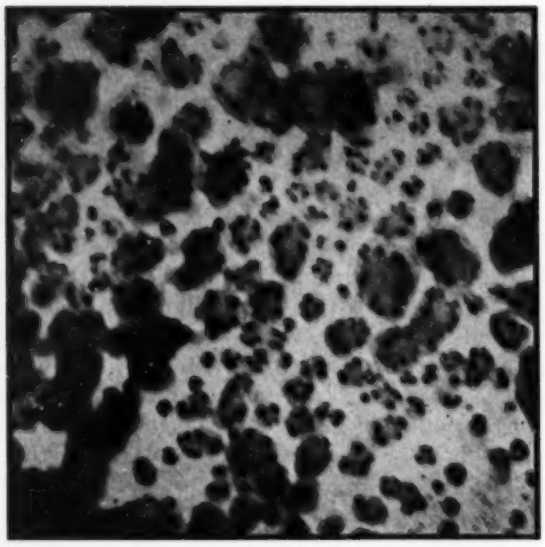


FIG. 6. Nebular structure obtained from the direct drying of a highly diluted emulsion polymer latex. ×14,000.

direct drying of a 1-ppm solution of polychlorostyrene from its dioxane solution (Fig. 7). In this latter case the dioxane may have contained some water, or picked up moisture during the drying of the drop. It is interesting to note, without necessarily drawing any inferences therefrom, that an astronomical nebula discovered at the Mt. Wilson Observatory was strikingly similar in appearance to the polymer systems under discussion.[34]

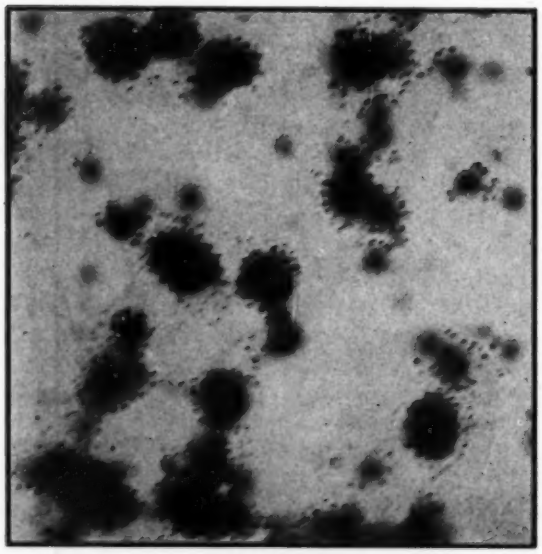


FIG. 7. Nebular pattern resulting when a 0.0001 percent solution of polychlorostyrene in dioxane was dried directly on the silica substrate. ×32,000.

PARTICLE SIZE DISTRIBUTION STUDIES

It was anticipated that if the solvent-precipitant technique was capable of isolating single polymer molecules, then with decreasing polymer concentration a point should finally be reached below which particle size became independent of solute concentration. The ultimate particle, supposedly a polymer molecule, should then be obtained. It is seen, from a study of Fig. 3, that at concentrations of 10 ppm and below, this stable plateau of particle size is reached to a good approximation. The nucleus disappears and the average diameter of the nebulae decreases below 10 ppm, but the particles on the periphery no longer change in size appreciably. Since coagulation and aggregation into micelles are both strongly dependent on concentration, it would be judged that both processes have ceased to operate at 10 ppm and below.

TABLE I. Observed particle size count for *p*-chlorostyrene.

0.1 ppm		10 ppm	
Diameter, d[a] A	No. of particles	Diameter, d[b] A	No. of particles
64	1	63	13
85	53	78	8
106	12	94	64
128	100	110	41
149	27	125	81
170	55	141	20
191	7	156	54
212	29	172	11
234	4	188	22
255	2	203	5
		219	4
		312	1
290 Particles $\bar{d}=141$A		324 Particles $\bar{d}=128$A	

[a] Obtained with binocular microscope.
[b] Obtained with spectrum reading magnifier.

Particle size distributions are presented in Table I on two fields each approximtely six microns square, the one at 10 ppm, the other at 0.1-ppm concentration of polymer. Only the small black particles connected with the nebular clusters were counted and measured. The procedure was as follows. The original plates, taken at a magnification of 8000, were enlarged to 32,000 photographically. Particle diameters were

[34] F. Zwicky, *Théodore von Kármán Anniversary Volume* (California Institute of Technology, 1941), Fig. 4 of Plate II. We are indebted to Dr. Walter S. Adams, Director of the Mt. Wilson Observatory, for negative reproduction of this nebula.

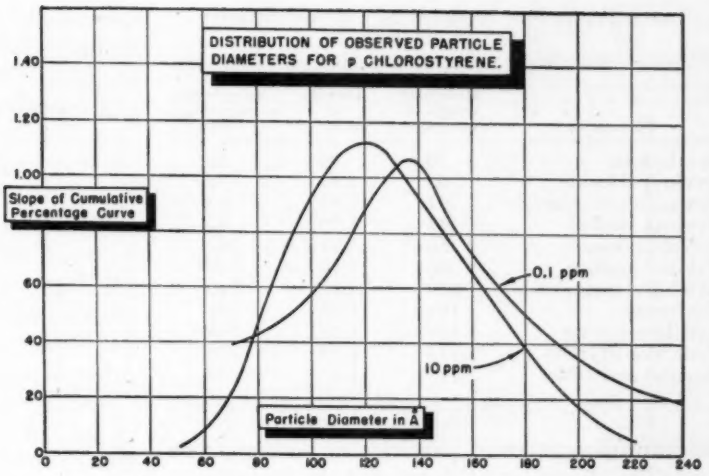


FIG. 8. Observed number distribution of particle diameters at two concentrations of polychlorostyrene. (Based on the data in Table I).

read from the paper prints, illuminated from behind. At first a binocular microscope with a scale in the eyepiece was used. Later, a Bausch and Lomb spectrum reading magnifier, with a 10-fold magnification, and a scale reading directly to 0.1 mm was employed. The image boundaries were diffuse and there was a tendency to read to the nearest 0.1 mm. This accounts for the oscillation in count of particles shown in Table I. The small counts at alternate numbers represented estimations to 0.05 mm. It is recognized that there are certain standard errors in photographic measurements of this type, such as shrinkage of the photographic emulsion and definition of the outer boundary of the image. There effects can be ignored at the present time in view of larger errors to be discussed later.

Cumulative percentage *versus* particle size curves were then constructed from the data in Table I and differentiated graphically to yield the smoothed-out particle size distribution curves of Fig. 8. It is seen that there is a slight trend toward larger particles at the lower concentration, and at the same time an indication of more small particles. This may imply that even at 10 ppm some of the larger molecules were involved in aggregation, for example, in forming the nucleus of the nebulae. Statistical fluctuations undoubtedly account for some of the difference.

Tables II and III present a more detailed study of this particle count data. The data are organized in terms of the number, n, of small particles found in each cluster; the number of clusters (in a field 6 microns square) containing a given number of particles; the approximate average diameter, φ, of the clusters, the average diameter, d, of the particles in each cluster; the average particle size, $\bar{d}$, for all clusters containing n particles; and finally the maximum and

TABLE II. Size studies on *p*-chlorostyrene—10 ppm.[a]

No. of particles in cluster, n	No. of such clusters	Cluster diameter, φ[b]	Average diameter of all particles in a given cluster, d	$\bar{d}$	d_{max}	d_{min}
10	1	1560	123	123	172	94
14	3	2080	123, 128, 132	128	187	62
15	1	2340	139	139	187	94
16	1	1560	123	123	156	94
17	1	2180	128	128	187	94
18	2	2180	124, 128	126	218	62
21	1	2340	141	141	218	62
22	3	2430	127, 129, 130	129	218	78
24	1	2800	130	130	187	62
25	1	2500	124	124	202	62
26	2	2500	113, 134	124	312	62

[a] All dimensions in angstrom units.
[b] Defined by Fig. 4.

TABLE III. Size studies on *p*-chlorostyrene—0.1 ppm.[a]

No. of particles in cluster, n	No. of such clusters	Cluster diameter, φ[b]	Average diameter of all particles in a given cluster, d	$\bar{d}$	d_{max}	d_{min}
10	1	1070	117	117	191	85
11	1	1070	121	121	170	85
13	3	1400	121, 122, 186	132	233	85
15	3	1300	119, 122, 136	125	212	85
16	2	1470	128, 154	141	170	64
17	2	1740	128, 179	153	254	85
18	1	1800	154	154	254	85
19	1	1800	149	149	212	85
20	4	1600	138, 140, 144, 159	146	212	85

[a] All dimensions in angstrom units.
[b] Defined by Fig. 4.

TABLE IV. Volume occupied by a monomer unit in a bulk polymer

Polymer	Molecular weight of monomer	Density of polymer	Volume per monomer unit A^3
Polyethylene	28	.92	50
Polyvinyl chloride	63	1.41	73
Polyvinylidene chloride	97	1.87	85
Polyvinyl alcohol	44	1.26	92
Polyisobutylene	56	.912	102
Polyvinyl acetate	86	1.19	119
Polymethyl methacrylate	100	1.18	140
Polystyrene	104	1.05	164
Polychlorostyrene	139	1.18	194
Polydichlorostyrene	173	1.39	206
Polyvinyl carbazole	193	1.2	256

minimum diameters of particles in a cluster are listed. The diameter, $\bar{d}$, increases approximately with the number of particles in the cluster. At 10 ppm the average diameter $\bar{d}$ is independent of cluster size, whereas at 0.1 ppm $\bar{d}$ increases with φ. At both concentrations the maximum and minimum particle diameters show no marked dependence on φ. While φ and n are both larger on the average at the higher concentration, yet values of $\bar{d}$, d_{max}, and d_{min} are approximately independent of concentration.

These two sets of data covering a 100-fold range in polymer concentration are believed to indicate that the particles observed are actual polymer molecules. The next section will point out, however, that molecular weights calculated from these particle diameters by several plausible assumptions give values which are too high by almost one order of magnitude.

RELATIONSHIP BETWEEN IMAGE DIAMETER AND MOLECULAR WEIGHT

Granting that the small particles of Table I, ranging in diameter from 60 to 300A, represent individual polymer molecules, the most important question then becomes the numerical relationship between measured image diameter and molecular weight of the particle. If one assumes that these particles have the same density as the bulk polymer, then molecular weight is simply proportional to diameter cubed. Table IV shows the volume required per monomer unit in the bulk polymer for a number of plastic materials. These values were calculated from the molecular weight of the monomer and the density of the polymer. It is evident that polychlorostyrene combines the advantages of a high electron scattering power and a large volume per monomer, thus leading to large sized particles. Polyethylene, on the contrary, would give rise to extremely small particles if bulk density prevailed. Polystyrene, with a density of 1.05 requires 164A³ per monomer unit, while polychlorostyrene with a gravity of 1.18 would correspond to 194A³ per monomer unit. Application of this latter value to the particles of Table I gives an abnormally high value for the average molecular weight, and a distribution curve which is much broader than that normally observed for high polymers. While these are not valid reasons against the assumption of bulk density, yet it seems worth while to investigate some other aspects of the problem.

Incidentally, Langmuir[35] has made some interesting calculations based on surface tensions and heats of evaporation of paraffins which lead him to believe that paraffin molecules in the vapor phase might exist as spheres.

A high polymer molecule in its most probable state can be considered, as was mentioned earlier, as a randomly-kinked chain which meanders through a volume of space proportional to the three-halves power of the number of carbon-carbon bonds in the chain. This volume has been defined by Kuhn[36] in a somewhat arbitrary manner as an ellipsoid whose root mean square axes have the ratio of 6:2.3:1, and whose cubical contents are given by

$$\text{Volume} = 0.83\gamma^{3/2}P^{3/2}. \qquad (3)$$

P = degree of polymerization; γ = number of chain links in a statistically independent unit. Such questions as free rotation about the carbon-carbon bonds, and steric hindrance from large substituent groups, as well as the space requirements of a physical chain in distinction to a mathematical chain must be considered. Kuhn therefore introduced a factor, γ, based on the assumption that free rotation occurs not at every carbon atom but at every γ carbon atom. Since γ will not be known in general, its presence thus constitutes either an uncertainty in computing molecular weights or a parameter which can be used to adjust or correct molecular weights obtained with the electron microscope. Kuhn further considered that steric effects might increase the exponent in Eq. (3) from 1.5 to 1.8 or 1.9.

Kuhn's prediction of a statistical volume proportional to a power of the molecular weight between 1.5 and 2.0 appears to be roughly ful-

[35] I. Langmuir, Proc. Roy. Soc. A170, 9 (1939).
[36] W. Kuhn, Kolloid Zeits. 68, 2 (1934); Zeits. f. angew. Chemie 49, 858 (1936).

filled for the particles of Table I. This may be seen from the plots prepared in Fig. 9 and based on data given in Table I. The observed particle frequency has been plotted against various functions of the measured particle diameter. Curve *A* corresponds to bulk density while Curve *B* refers to the Kuhn random coil according to Eq. (3). Curve *C* corresponds to Eq. (3) when the exponent of *P* is 2 instead of 1.5. Judging from molecular number distribution curves which have been obtained on polystyrene,[37,38] Curves *B* and *C* appear much more plausible than the bulk density assumption.

However, within the limits of observation, the particles are circular in cross section, and not elongated as Kuhn's theory predicts. The most extreme departure from circular images observed thus far was on a chlorostyrene sample polymerized at 60°C to a weight average molecular weight of 680,000 (by light scattering) and precipitated from a 0.1 ppm benzene solution according to a technique which will be mentioned later. Out of 314 particles measured, only 20 particles showed a noticeable lack of symmetry. Table V summarizes the observations.

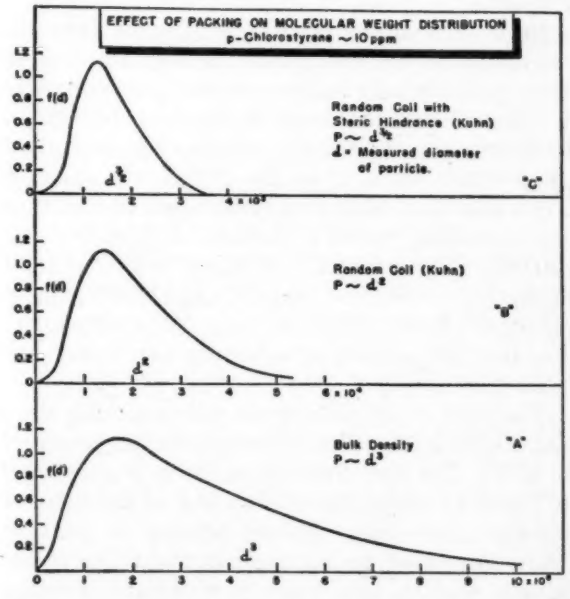


FIG. 9. Molecular weight number distribution curves based on several assumptions. The ordinates represent the observed frequency distribution of particle sizes listed in Table I. The abscissae correspond to various functions of the measured particle diameters.

[37] H. Mark and R. Raff, *High Polymeric Reactions* (Interscience Publishers, Inc., New York, 1941), pp. 47–59, p. 197.

[38] G. V. Schulz and A. Dinglinger, Zeits. f. physik. Chemie **B43**, 47 (1939).

TABLE V. Asymmetry of some measured polychlorostyrene particles.

Axis ratio	No. of particles observed
1.0	294
1.5–2.0	3
2.0–2.5	6
2.5–3.0	6
3.0–3.5	2
3.5–4.0	1
4.0–4.5	1
4.5–5.0	1
	314

The uncertainty of 0.05 mm in measuring diameters could mean that many particles were out of round by this amount. The average diameter particle in this series might then have had an axis ratio of 1.15 to 1. In several cases listed in the table, there was a distinct suggestion of two or more particles in close contact, rather than of one highly elongated particle. While Kuhn's general treatment of a random coil appears correct, yet the arbitrary model he chose for computing volumes does not seem applicable in the present case.

A more recent and more pertinent treatment of the problem of a random coil is that given by Daniels.[39] He proposes the question: What is the most probable size of a box which will just completely enclose a random coil? He assumes that a given point of space can be occupied any number of times by different parts of the chain, and also that each succeeding link is in an entirely random direction in space. He finds for the most probable extent, ρ_{max}, of such a coil in one dimension

$$\rho_{max} = 1.32a\sqrt{N}, \tag{4}$$

where "a" is the length of a link and N is the number of links in the chain. This is a limiting form for large values of N. Since the polymerization degree, P, is approximately $N/2$, we have

$$\rho_{max} = 1.87a\sqrt{P}. \tag{5}$$

Identifying ρ with the measured particle diameter, this equation predicts a volume proportional to $P^{3/2}$, or a molecular weight proportional to $(\rho_{max})^2$. Daniel's treatment, as is inevitable from its statistical nature, likewise predicts a distribution in particle diameters ranging from roughly ½ to 2.5 times the most probable extent. He shows, moreover, that for large P the extent in one direction is independent of the extent in

[39] H. E. Daniels, Proc. Camb. Phil. Soc. **37**, 244 (1941).

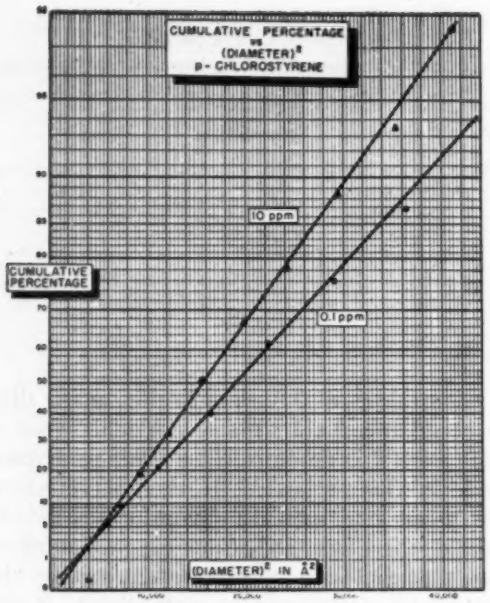


FIG. 10. Cumulative number percentage distribution curve *versus* the square of particle diameter for the data of Table I. The abscissae are assumed to be proportional to the molecular weights of the individual particles.

the other two perpendicular directions. However, the sharp distribution function which he obtains suggests that most of the images will be very close to circular in shape.

If it is assumed that the particles discussed in Table I follow the law of a random coil, then their molecular weights will be proportional to the square of their respective diameters. Figure 10 shows plots of cumulative number percentage against diameter square. A special type of graph paper designed to yield straight lines for molecular weight distribution data of the type observed on polystyrene was employed here.[40] The fact that straight lines are obtained suggests first that the random coil concept is reasonable, and also that the number distribution equation of molecular weights is of the form

$$N(P)dP = \text{constant } P^b\alpha^P, \tag{6}$$

while the weight distribution equation is of the type

$$W(P)dP = \text{constant } P^{b+1}\alpha^P, \tag{7}$$

where α is a parameter lying between .99 and 1.0 and P is again the degree of polymerization. b is a parameter whose value usually lies between 0 and 4,[37,40] and which has a value of 2 for the data of Fig. 10. The molecular weight of a polymer molecular is of course $P\,Mo$ where Mo is the molecular weight of a monomer unit. $N(P)$ is the fractional number, and $W(P)$ is the fractional weight of chains having polymerization degrees between P and P plus dP.[38] The weight-average molecular weight, $\bar{M}w$, and the number average molecular weight, $\bar{M}n$, are given by

$$\bar{M}w = Mo(b+2)/(1-\alpha),$$
$$\bar{M}n = Mo(b+1)/(1-\alpha). \tag{8}$$

In our use of Daniel's one-dimensional case, the length of the link, "a," will correspond to the projected length of a carbon-carbon bond. If these bonds were distributed at random in space, this projected length would be $lo/\sqrt{3}$ where lo is the carbon-carbon bond length, namely 1.54A. However, for tetrahedral angles Kuhn showed that the projected length is $lo(2/3)^{\frac{1}{2}}$ and this value should be substituted for "a" in Eq. (5). If, furthermore, the factor γ is introduced to provide for lack of free rotation, there results

$$\rho_{max} = 2.32\gamma^{\frac{1}{2}}P^{\frac{1}{2}}; \tag{9}$$

or

$$\rho_{max}^2 = 5.38\gamma P. \tag{10}$$

From the nature of the distribution function predicted by the plots[40] of Fig. 10, the weight average molecular weight would correspond to a value of the diameter squared of 24,000 for 10 ppm or to 28,000 for 0.1 ppm. Substituting these values for ρ_{max}^2 in Eq. (9) with γ unity, P is calculated as 4400 and 5200, respectively, thus corresponding to $\bar{M}w$ values of 610,000 and 720,000, respectively. Compared with the light scattering molecular weight of 140,000, these values are 4 to 5 times too high. We assume here that the light scattering value is a weight average molecular weight.[41]

The only other sample on which similar data is available is the chlorostyrene polymer prepared at 60°C. The size distribution data is presented in Table VI while Fig. 11 is a plot of cumulative number percentage against square of particle diameter. Again an excellent straight line is obtained, leading this time to a weight average molecular weight of 3,000,000 which is 4.4 times the light scattering value of 680,000. b in this case has a value of unity.

In each case the electron microscope method gave weight average molecular weights four to five times higher than those obtained by light

[40] R. F. Boyer, manuscript in preparation.

[41] R. S. Spencer, The Dow Chemical Company, unpublished calculations.

scattering. There are several possibilities which may account for such a discrepancy.

1. Aggregates of several molecules are involved. This would appear to be ruled out by the relative constancy of molecular weight with concentration. In fact, molecular weight increased from 610,000 to 720,000 on decreasing the concentration of polymer from 10 to 0.1 ppm. This is opposite to what might be expected.

2. The particles as deposited on the silica screen are still swollen with solvent which is removed so rapidly in the electron microscope that the particles cannot shrink to an appropriate dry volume. Because of the large proportion of precipitating agent used in these experiments, we would expect this swelling effect to be small.

3. The particles are spherical and of the correct diameter when suspended in the solvent mixture, but are flattened by attractive forces between silica screen and polymer.

4. The factor γ in Eq. (9) is 4 to 5 instead of unity as was used in calculating molecular weights. This would mean that each four to five carbon-carbon links act as a statistically independent unit. Such a value for γ is not unreasonable. In fact, Kuhn has recently calculated values for γ from streaming double refraction measurements on polymer solutions. For styrene polymerized at 20°C γ is 14 in toluene, while for polystyrene made at 135°C γ is 5, again in toluene.[42]

It does not seem possible to decide conclusively at present which, if any of these factors, is operating to produce high values of molecule weight. Various experiments suggest themselves as means of clarifying some of the questions. However, Fig. 12, showing plots of expected particle diameter as a function of polymerization degree, is instructive. The lower heavy curves are based on the assumption of bulk density for five different polymers. The dotted lines represent the theoretical values for coiled chains according to Eq. (9), and under the assumption of different values for γ.

TABLE VI. Observed particle size distribution for *p*-chlorostyrene polymer prepared at 60°C.

Diameter A	No. of particles	Diameter A	No. of particles
48	1	306	3
64	2	322	26
80	1	338	5
96	4	354	14
113	3	370	11
129	10	386	14
145	1	402	3
161	27	434	1
177	7	450	9
193	29	466	1
208	5	482	4
225	41	498	1
242	13	514	4
258	23	546	1
274	5	645	1
290	24		

Total no. of particles—294
Average diameter —267A

Note: The 20 elongated particles also observed in this series are reported separately in Table V.

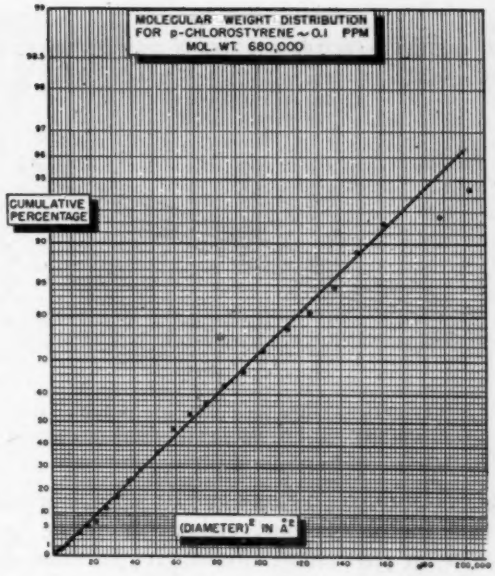


FIG. 11. Cumulative number percentage distribution curve *versus* square of particle diameter for the high molecular weight polychlorostyrene sample of Table VI.

One notes that for γ of unity the volume of a random coil becomes less (and hence its density becomes greater) than that of bulk polychlorostyrene for all polymerization degrees less than 1000. Now it does not seem possible that the solvent-precipitant system used could compress a polymer chain into a greater density than that of bulk polymer. Yet the cumulative percentage curves of Figs. 10 and 11 suggest that random coils and not bulk density are involved. If, however, γ has a value somewhere near 4, most of this discrepancy in density is removed, as can be seen on the graph. At the same time a γ of 4 would provide for most of the discrepancy between light scattering and electron microscope molecular weights. This value of γ might reflect the fact that the large size of the chlorobenzene groups attached to the hydrocarbon skeleton, as well as repulsion of chlorine atoms for each other provide enough steric hindrance to prevent the polymer chain from coiling up as tightly as a mathematical chain might. This argument is not offered as a proof that single molecules are involved in this study and that the molecular

[42] W. Kuhn and H. Kuhn, Helv. Chim. Acta **26**, 1394 (1943).

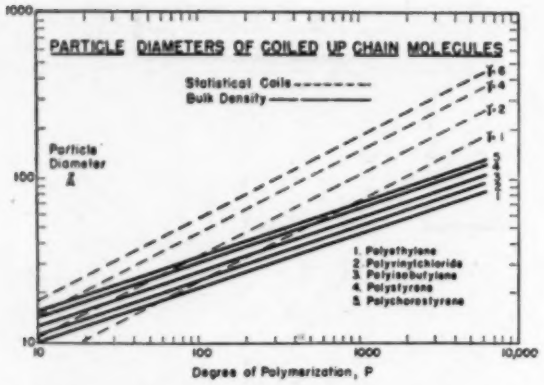


FIG. 12. Theoretical diameters of single polymer particles as a function of the degree of polymerization.

weights obtained on them can be safely corrected by this factor of four. Other polymers, polyethylene in particular, might give random coils with γ of unity without violating the bulk density requirement appreciably. Certainly γ would be expected to differ for different types of polymers. It might be thought that the coiled up particles could follow two laws of particle diameter *versus* molecular weight: The random coil law for P greater than 1000, the bulk density law for smaller molecules. However, the cumulative percentage plots in Figs. 10 and 11 suggest only the random coil. A case will be mentioned later where this dual type of behavior seems indicated.

It thus seems that the problem of resolving single polymer molecules into particles giving circular images in the microscope can be solved experimentally in certain cases. The complementary problem of converting image diameters to molecular weights is in a far less satisfactory state. However, if the electron microscope values of molecular weight are relatively correct within themselves, the distribution curve will be correct as to shape and extent. The absolute values can then be provided by some independent method such as light scattering.

As regards the shape of the distribution curve, it was recognized that even with a single molecular weight species present, there should be a distribution in image diameters of the type predicted by Daniels. With a heterogeneous polymer each molecular weight species would give a distribution in image sizes. Hence the observed particle size distribution curve should involve the superposition of both effects. A theory accounting for this has been developed (see Appendix) and applied to the experimental data. It turns out that the experimental curves are sharper than the theory would predict. Figure 13 is a plot of the observed particle size distribution data given in Table VI. It is compared with Daniel's statistical distribution for a single molecular weight material, and our theoretical distribution, Eq. (16) in the appendix, for $b=1$. The two theoretical curves were adjusted to have a maximum at 220A as required by the experimental data.

We realize that there is a logical contradiction in what we are trying to do here. At the moment we do not see how to resolve the conflict. We say, for example, that the straight line plots in Figs. 10 and 11 imply random coils, and a size distribution function as given by Eq. (6). This would be true if each molecular weight species resulted in only one particle diameter. It might appear that this problem would be simplified if we worked with narrow fractions. However, a polymer fraction will contain a distribution of molecular weights. A polystyrene fraction subsequently studied in the ultracentrifuge showed a range of molecular weights from 30,000 to 150,000.[43] It will apparently be necessary to work with polymers whose molecular weight distributions are accurately known by other methods before these details can be investigated in full. In the meanwhile, it is curious that Figs. 10 and 11 represent exactly the type of distribution data that have been obtained on polystyrene.[37,38,40] A possible explanation is given in the following paragraph.

Orr[7] has recently pointed out that a polymer molecule dissolved in a solvent which gives strong heat interaction with the polymer cannot behave in a truly random fashion. The solvent tends to promote polymer-polymer contacts. This, of course, was the philosophy behind our adding a precipitant to the polymer solution

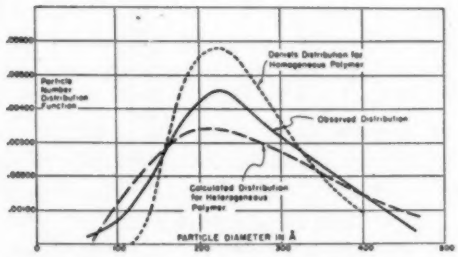


FIG. 13. Theoretical and observed particle size distribution curves for the data of Table VI.

[43] Figure 9 on page 55 and the table on page 61 of reference 37.

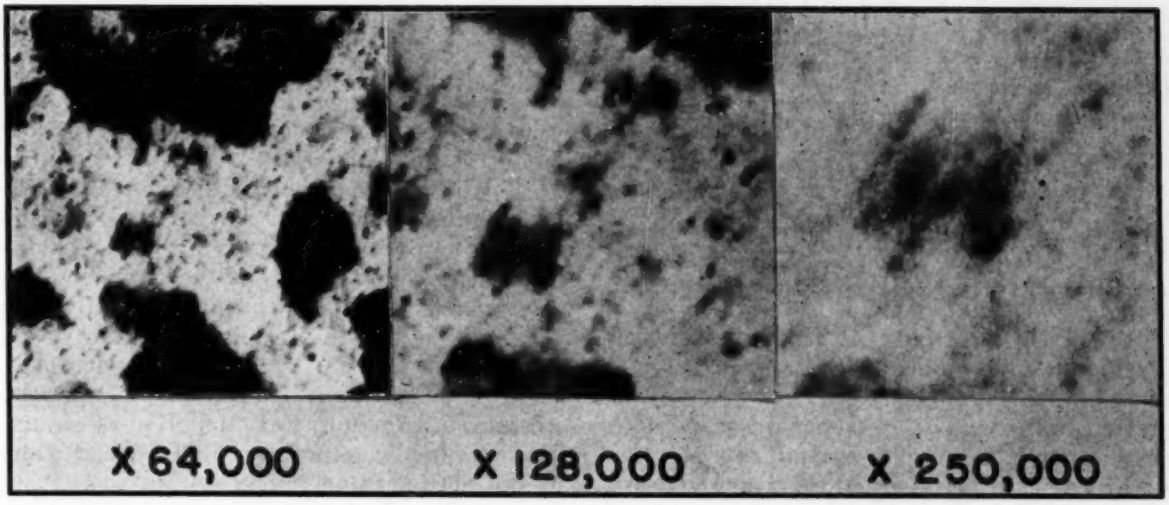


FIG. 14. Pattern obtained by adding ice to a 0.0001 percent solution of polychlorostyrene in dioxane. This polymer was the same one used for Fig. 3.

so as to coil up an otherwise extended chain molecule. In view of Orr's suggestion, it would then seem that any molecular chain which, in a state of complete randomness preferred to be less tightly coiled than the average chain, would be forced by the poor solvent to assume a more nearly average configuration. On the other hand, geometrical problems would place a lower limit on the tightness of packing. The interplay of these two factors might account both for the lack of asymmetry in the electron microscope images and for the sharpness of the observed number distribution curves, as given in Fig. 13. Some question remains as to whether the volume of such a "random" coil would still depend on the three-halves power of the number of links in the chain.

MISCELLANEOUS EXPERIMENTS

Several attempts were made to deposit polymers directly from dilute solutions. The work with dioxane was shown in Fig. 7 with the suggestion that traces of water may have caused some coiling up of particles. Direct drying from a 1-ppm benzene solution gave what appeared to be a continuous film with here and there some large irregular masses and a few elongated objects (about 50×400A). One ppm of polyethylene dissolved in benzene at 80°C and allowed to cool and evaporate on a silica screen showed mainly agglomerates and an occasional circular particle.

One particularly interesting result occurred when a pure piece of frozen distilled water was placed in a 1-ppm dioxane solution of chlorostyrene. The work of Staudinger and Heuer[15] had indicated that water was an extremely effective precipitant for dioxane solutions of polystyrene, as was mentioned earlier. The results obtained are presented in Fig. 14 showing magnifications of 64,000, 128,000, and 250,000. In addition to large irregular masses there can also be seen, especially at the highest magnification, some extremely small particles. A number of images between 15 and 50A have been noted. These are smaller than the value of 40A given by von Ardenne for organic molecules.

Eighty-nine particles selected at random on a photograph at 64,000 magnification were measured. They showed a range in diameter from 31 to 156A, an average diameter of 73.5A. Both the range and the average diameter are considerably smaller than that found on this same polymer using the benzene-propanol system. A molecular weight distribution curve obtained on this admittedly small number of observations behaved as if bulk density were being followed up to a polymerization degree of 900, and a random coil law for higher values. The weight average molecular weight calculated on such a basis was 210,000 as compared with the light scattering value of 140,000. While the dioxane-ice system resolved a number of the lower molecular weight particles, it may have forced the large molecules into aggregates so that the distribution curve and the relatively low molecular weight obtained may not be representative of the true distribution.

TABLE VII. Some properties of a solution containing polystyrene molecules of molecular weight 100,000.

Concentration of polymer ppm	No. of particle per cc $\times 10^{-13}$	Edge of a cube containing one polymer molecule A	Coagulation time in milliseconds
100	606	1170	.9
10	60.6	2520	9.0
1	6.06	5430	90
0.1	.606	11700	900
0.01	.0606	25200	9000

It is readily apparent that type of solvent and precipitant can influence the results obtained considerably. This is in agreement with what has previously been reported from viscosity studies on high polymers. A great amount of systematic work must still be done before proper solvents and precipitants can be selected. It was noted that some reagents, namely benzene, and water were relatively easy to purify by a simple flash distillation. Other reagents such as methyl ethyl ketone, dioxane, and *n*-propanol could not be purified even after repeated distillation. Purity is used here in the sense of running a blank by placing a drop of the liquid on a silica screen and examining it in the electron microscope. When working with polymer concentrations of 1 ppm it is evident that foreign material in the solvents used must be rigidly excluded. While *n*-propanol and methyl ethyl ketone could be made dust free, yet they produced amorphous deposits, presumably polymeric in nature. This effect was especially noticeable if the reagents had aged for a month or so in contact with air.

EFFECT OF ELECTRIC CHARGE

There is one further point of considerable importance which became apparent from these early studies. It concerns the role of electric charges in stabilizing the existence of small particles. The original sample of polychlorostyrene (molecular weight 140,000) was resolved into seven fractions by a precipitation technique, in order to observe the performance of fractions in the microscope. Every effort to obtain small spherical particles from these fractions met with complete failure. The field of the microscope was covered with large irregular agglomerates in every case.

A review of the situation suggested that electrolytes or other polar groups which would retard coagulation might have been playing an unsuspected part in the original experiments. This first polymer sample studied was at least two years old, and may have been made from slightly impure monomer, or may have oxidized slowly from contact with air. Moreover, the solvents and precipitants used in the early work were only moderately clean and of the type commonly found around a research laboratory. Purification of this polymer by precipitation, and purification of the solvents by repeated distillation presumably gave a system so free from electrical charges that the polymer molecules coagulated either during or shortly after the addition of the precipitant, or during the drying stage. At any rate it was evident that whatever success attended the early experiments arose from a fortuitous combination of circumstances which were not then realized and which are not yet fully understood.

Some appreciation of the inherent difficulties involved can be gained from the numerical values given in Table VII. This table shows the volume available for a single polymer molecule of average molecular weight 100,000 at several concentrations, the number of particles per cc, and also the coagulation time calculated according to von Smoluchowski, that is, assuming no charge on the particles.[44] Aging studies on the 10-ppm system indicated stability for at least 72 hours after addition of the precipitant; but coagulation between 72 and 168 hours. It is thus evident that the particles must have been highly charged.

Experiments were, therefore, tried in which a trace of HCl was added to the *n*-propanol used for precipitation. One drop of concentrated HCl in 50 cc of the alcohol made a profound change in the character of the images, as can be seen in Fig. 15. This is the chlorostyrene sample of 680,000 average molecular weight whose size distribution was given in Tables V and VI. Figure 16 shows the pattern obtained on fraction 4 of the polymer of Table I (mol. wt. of fraction 110,000) by using HCl. A new type of structure resembling a honeycomb is seen here. Each cell contains one or two particles whose size is again of the correct order of magnitude. Three hundred and ninety-seven particles ranging in diameter from 63 to 240A were counted in one field. The average diameter was 130A. The distribution curve showed fewer small and fewer large particles than did the parent sample from which the fraction was obtained. Thus the distribution curve was slightly sharper than that for the

[44] H. R. Kruyt and H. S. van Klooster, *Colloids* (John Wiley and Sons, New York, 1930), p. 117ff.

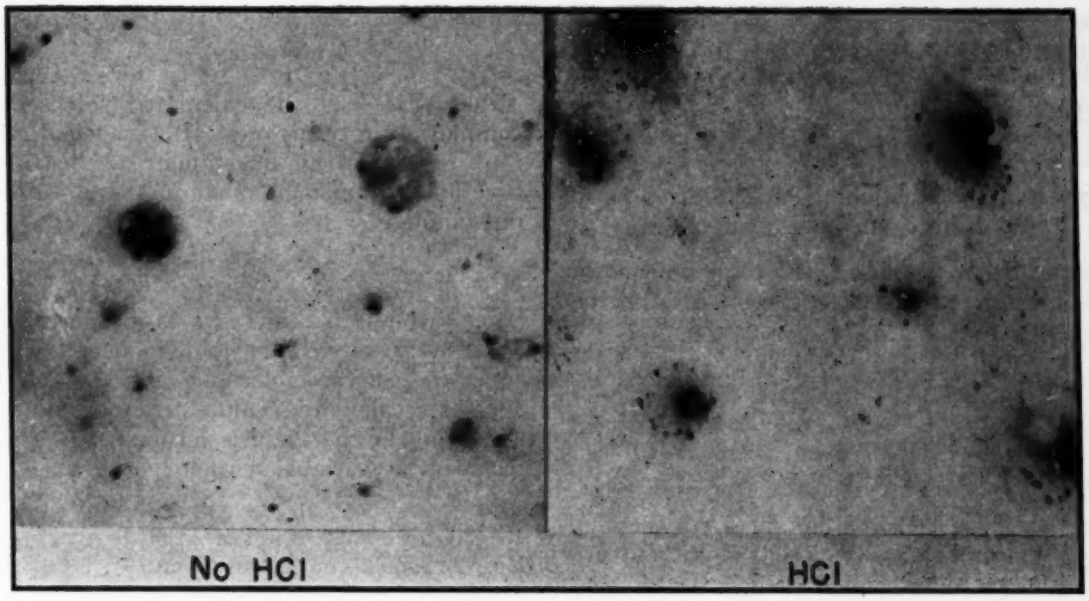


FIG. 15. Effect of traces of HCl on the structure obtained by adding *n*-propanol to a 0.00001 percent benzene solution of polychlorostyrene. The data for Table VI were obtained from the above plate with HCl. ×16,000.

polymer. However, the distribution data did not follow a Gaussian law, nor did it give linear plots on the type of graph paper used for Figs. 10 and 11.

Time has not permitted a more thorough investigation of this aspect of the problem. It seems fairly definite that electric charges must either be present or be added to the solvent system in order to resolve single particles. HCl is likely not the best type of addition agent. It may be that a substance like stearic acid or sodium stearate would produce more satisfactory results.

One final experiment relating to coagulation might be mentioned. It was thought that rate of addition of the precipitating agent to the high polymer solution might be critical Two parallel experiments were carried out. In one case that precipitant was added as rapidly as possible. In the second trial the precipitant was added dropwise along the side of the beaker, with continuous stirring, over a half hour period. The electron micrographs obtained by these two treatments were very similar. If anything, the rapid addition gave somewhat sharper particles.

ACKNOWLEDGMENT

This material was first presented in part before a joint meeting of the American Physical Society and the Electron Microscope Society in New York City, January 11–15, 1944; and in more detail before the Gibson Island High Polymer Conference, June 26, 1944. We are indebted to a number of individuals who offered comments. We should particularly like to thank R. Simha, L. A. Matheson, H. Mark, and P. M. Doty with whom we reviewed various aspects of the problem in more detail. In presenting this material at the present time it is recognized that the

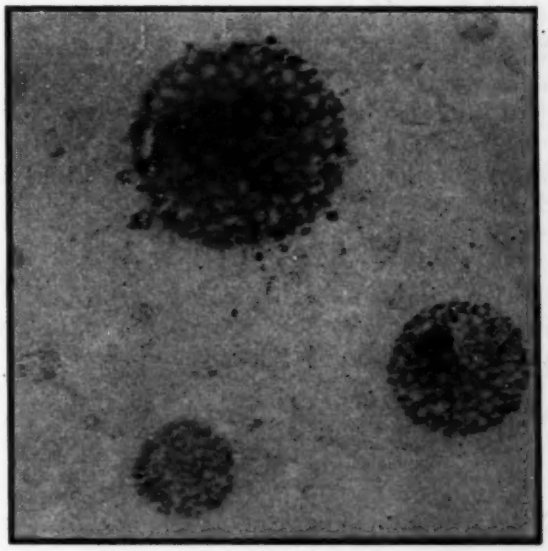


FIG. 16. Honeycomb structure obtained through the use of HCl—benzene—propanol on a fraction of polychlorostyrene. There are one or more particles in each of the cells of the honeycomb. ×33,000.

problem is far from completed. However, the results obtained and some of the difficulties which have appeared are believed to be of general interest.

APPENDIX

Expected Distribution of Particle Sizes for a Heterogeneous Polymer

According to Daniels,[39] the probability p_N of finding a random coil of linear extent ρ is given by

$$p_N(\rho)=(8/(2\pi Na^2)^{\frac{1}{2}})$$

$$\times\sum_{k=1}^{\infty}(-1)^{k-1}k^2\exp(-k^2\rho^2/2Na^2), \quad (11)$$

where a is the length of a link, and N is the number of such links. $N\sim 2P$ where P is the degree of polymerization, while $a=lo(2/3)^{\frac{1}{2}}$ as given earlier. For a polymer containing a distribution of molecular weights of the type indicated by Fig. 11, or Eqs. (6) and (7) with $b=1$, the probability $N(P)$ of having a molecule of polymerization degree between P and $P+dP$ is:

$$N(P)dP=(1-\alpha)^2P\alpha^P dP, \quad (12)$$

$$=(1-\alpha)^2P\exp(-B^2P)dP, \quad (13)$$

where $B^2=\log_e\alpha\sim-(1-\alpha)$. The expected size distribution of particles to be observed in the electron microscope should be given by multiplying Eqs. (11) and (13) and integrating. This leads to integrals of the form

$$\int_0^{\infty}P^{\frac{1}{2}}\exp(-B^2P-q^2/P)dP, \quad (14)$$

whose solution has been given by Bierens de Hahn[45] as:

$$[(\pi)^{\frac{1}{2}}/2B^3](1+2Bq)\exp(-2Bq). \quad (15)$$

Carrying through the indicated operations, we arrive at the following expression for the frequency distribution of particle sizes:

$$F(\rho)=\frac{2(1-\alpha)^{\frac{1}{2}}}{a}$$

$$\times\sum_{k=1}^{\infty}(-1)^{k-1}k^2[1+kw]\exp(-kw), \quad (16)$$

where

$$w=(1-\alpha)^{\frac{1}{2}}\rho/a. \quad (17)$$

Table VIII lists the calculated values of the summation in terms of w. It is evident that Eq. (16) represents an unsymmetrical distribution function with a single maximum. Eq. (16) has a maximum at

$$\rho_{max}=2.305a/(1-\alpha)^{\frac{1}{2}}. \quad (18)$$

Since, from Eq. (8), Mw is proportional to $1/(1-\alpha)$; it is seen that ρ_{max} increases with the squre root of the molecular weight. Table VIII also lists calculated values for the summation appearing in the Daniels formula, Eq. (11), in terms of the variable u, where

$$u=\rho^2/4Pa^2. \quad (19)$$

Eq. (11) has a maximum given by

$$\rho_{max}^2=3.632Pa^2. \quad (20)$$

The equations just presented, and the numerical values in Table VIII, were used to compute the theoretical curves of Fig. 13. Thus far we have considered only the size distribution for

TABLE VIII. Numerical values for the summations appearing in Eqs. (16) and (11).

w	$\sum_{k=1}^{\infty}f(w)$	u	$\sum_{k=1}^{\infty}f(u)$
0.8	.0177	0.3	.015
1.0	.0470	0.4	.081
1.2	.0796	0.5	.172
1.4	.1070	0.6	.237
1.6	.1274	0.7	.274
1.8	.1464	0.8	.293
2.0	.1562	0.9	.299
2.2	.1634	1.0	.295
2.4	.1626	1.1	.284
2.6	.1579	1.2	.268
2.8	.1497	1.4	.232
3.0	.1393	1.6	.195
3.2	.1294	1.8	.162
3.4	.1161	2.0	.134
3.6	.1037	2.2	.111
3.8	.0918	2.4	.0907
4.0	.0795	2.6	.0743
4.2	.0698	2.8	.0608
4.4	.0607	3.0	.0498
4.6	.0525		
4.8	.0448		
5.0	.0385		
5.2	.0342		
5.4	.0289		
5.6	.0244		
5.8	.0206		
6.0	.0173		
6.4	.0123		
7.2	.00613		

[45] D. Bierens de Hahn, *Nouvelles Tables D'Intégrales Définies* (G. E. Stechert & Company, New York, 1939), Table 98, formula 4.

$b=1$. In general

$$N(P)dP=\frac{P^b\alpha^P}{b!}(1-\alpha)^{b+1}. \qquad (21)$$

In a manner similar to that used previously, the expected frequency distribution of particle sizes can be found.[46] The expression obtained is rather awkward to handle and no numerical calculations have been made with it.

[46] We are indebted to R. S. Spencer for this calculation. Table 98, formula 5 of reference 45 was used for the integration.

Conversion Diagrams for Triode Tube Mixers

HARRY STOCKMAN
Cruft Laboratory, Harvard University, Cambridge, Massachusetts
(Received June 28, 1945)

The theory of the frequency conversion diagram has previously been presented in this journal. The research work is now extended to include triode tubes, and it is shown in the following that the conversion diagram provides a suitable method of treatment of non-linear mixers with more than two electrodes. Actually, the greater the number of variables, and the greater the difficulties in a mathematical analysis, the more important becomes this method of obtaining frequency conversion data. A brief description of the measurement equipment is given in the text and typical diagrams and derived curves are shown.

As the Conversion Diagram method is already described in the literature,[1] the following discussion will be restricted to practical measurements on triode tubes as non-linear mixers. Only short wave measurements are described in this paper, but the information given about the measurement equipment brings out the fundamental principles, upon which the design of equipment for higher frequencies must be based. Formulas for the given conversion diagrams can be obtained in analogy with the similar treatment for diode tube and crystal mixers.[2]

While the triode tube is of very limited interest for medium-wave receivers, it has found considerable use in centimeter-wave receivers. As grounded-grid amplifier[3] the triode tube provides a practical circuit for stable high frequency amplification, as the tendencies for oscillations are almost entirely eliminated. The following discussion is merely part of an investigation of the possibilities of the grounded-grid amplifier as non-linear mixer. As only "zero" impedance signal-wave and oscillator-wave sources are considered in the following, the results apply as well, in principle, to a neutralized grounded cathode circuit. The discussion is primarily given to stress the possibilities of the conversion diagram technique and does not intend to provide a critical analysis of the behavior of a particular type of mixer tube.

THE MEASUREMENT EQUIPMENT

The experimental circuit is shown in Fig. 1, where A and B are the short wave signal generators for the signal wave of frequency $A/2\pi$ and the local-oscillator wave of frequency $B/2\pi$. The local oscillator actually employs a master oscillator and a following power amplifier, from which the output is obtained via a single-turn secondary coil as is indicated in Fig. 1. The triode tube used for these particular measurements was a "955" with the heater separated from ground by means of high frequency chokes and connected to the cathode by means of the coupling capacitor C_1. Cathode bias voltage is provided by the by-passed biasing resistor R_0 and can be extended by means of an additional

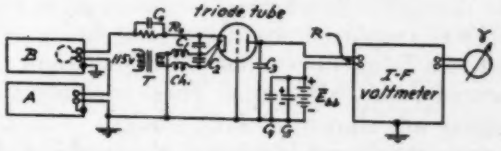


FIG. 1. Circuit diagram for the experimental measurement set-up. A grounded-grid circuit is employed with both heater and cathode at radiofrequency potential.

[1] Harry Stockman, "A Treatment of Non-Linear Devices Based upon the Theory of Related Linear Functions," J. App. Phys. **14**, 645–658 (1943).
[2] Harry Stockman, "Use of Frequency Conversion Diagrams," Proc. I.R.E. **32**, 679–686 (1944); "UHF Converter Analysis," Electronics **18**, 140–143 (1945); and Proc. National Electronics Conference, Chicago, **1**, 242–252 (October 5–7, 1944).
[3] Nils Erik Lindenblad, U. S. patents 2,136,448 and 2,151,781, filed 1932, patented 1938, respectively 1939.

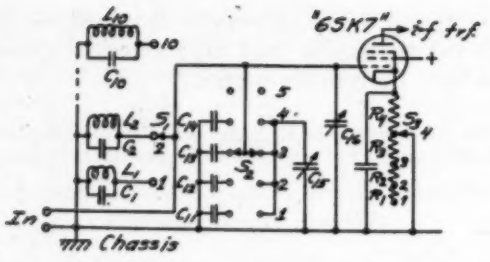


FIG. 2. The input circuit of the i-f voltmeter, showing plug-in coil-units L_1C_1–$L_{10}C_{10}$ with switching and tuning arrangements. S_3 is a sensitivity control.

bias battery. Plate bias voltage is obtained from a properly by-passed direct voltage source. The proper intermediate-frequency load impedance is provided by a specially designed i-f voltmeter, the indicating instrument of which is shown to the right in Fig. 1. This i-f voltmeter is calibrated in input intermediate-frequency voltage, and when the input impedance R is known the intermediate-frequency current and power can be calculated.

The photograph in Fig. 3 shows the most essential parts of the measurement set-up. The device in front is the mixer unit with the vertical filament chokes clearly visible. The coaxial cable coming in at a 45° angle carries the output from the A-wave source. The panel behind the mixer unit belongs to the local oscillator power amplifier, and the instrument on top of this amplifier is a vacuum tube voltmeter for the measurement of the power-amplifier output voltage. The probe to this voltmeter is visible in the lower left corner of the figure. The i-f voltmeter with its indicating instrument is shown to the right.

This brief discussion is sufficient to give a general idea of how the measurement equipment can be arranged, but further information is needed regarding the i-f voltmeter, as the results obtainable with the conversion diagram method depend very much upon the quality of this instrument. The input circuit to the i-f voltmeter is shown in Fig. 2. The input impedance R can be given ten different values, yielding a family of load-lines in the conversion diagram to be obtained. A scheme of exchangeable coil-units was employed, each coil-unit being self-tuned to a frequency somewhat higher than the chosen i-f value, 450 kc.p.s. These exchangeable coil-units are shown as L_1C_1, L_2C_2, $\cdots$ $L_{10}C_{10}$. The proper coil-unit is selected by means of the switch S_1, and tuning to zero phase angle is made possible by means of the fixed and variable capacitors C_{11}–C_{16} and the additional switch S_2. It should be noted that the input impedance value R is measured under conditions of zero phase angle, and that the higher impedances require careful adjustment of the vernier C_{16}. Figure 2 also shows the first i-f tube and a bias resistor switch S_3 for sensitivity control. The various controls are spread out over the panel as is shown in Fig. 3, the last knob to the right being the zero adjustment for the indicating galvanometer.

It is sometimes desirable that a particular region of a conversion diagram is carefully investigated. For such an investigation special plug-in coil-units are used, representing a group of closely-spaced load-lines in the particular region of interest. Wherever any two load lines are too much separated, a new coil-unit can be inserted to provide an in-between line. The slope of a particular load line may be adjusted by means of a changed L/C-ratio, or by changing the value of a shunting resistor attached to the coil-unit.

A triode tube mixer has higher output impedance than a diode tube or crystal mixer, and it is desirable that the bottom load line in the conversion diagram represents an R-value of several hundred thousand ohms. As is shown by Fig. 4 the writer used a maximum value of $R=160{,}000$ ohms, and a considerably higher value is difficult to obtain. While low impedance coils can be given small dimensions, high impedance coils of sufficiently high Q-value must have large dimensions and must be provided with large shielding cans, if shielding is required. Litz-wire with the maximum possible number of strands is desirable. Although the impedance of

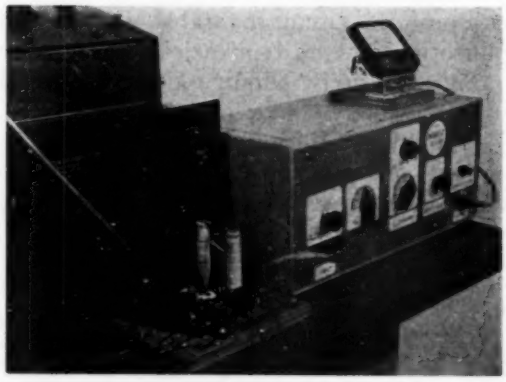


FIG. 3. Photograph of the measurement set-up showing the mixer unit in front with the oscillator power amplifier behind and the i-f voltmeter to the right.

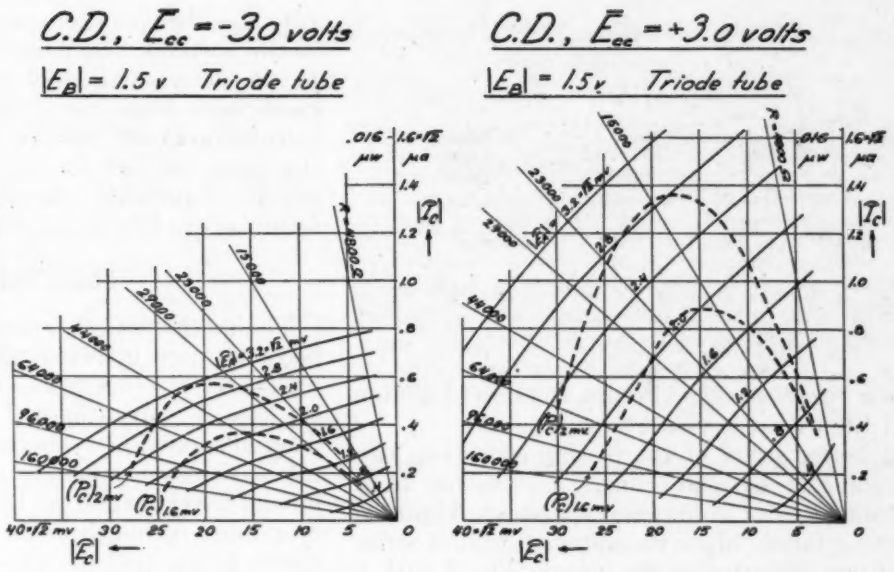


FIG. 4. Frequency conversion diagrams for different values of battery bias voltage $\bar{E}_{cc}$ with additional automatic bias obtained via a 5000 ohms cathode resistor R_0. The oscillator voltage is $E_B = 1.5$v r.m.s.

a "scientific" litz-wire self-tuned coil may approach the value one megohm, the practical designer must be satisfied with a much lower value primarily because of the unavoidable shunt capacitance provided by the capacitor C_3 between plate and ground in Fig. 1, plus stray capacitance, winding capacitance, etc. This shunt-capacitance is necessary to provide a by-pass for signal and oscillator currents. One effect of insufficient by-pass is an amplification of the order of one for signal and oscillator voltages, so that these voltages are impressed upon the input of the i-f amplifier. The oscillator voltage may then override the first i-f tube, and there are also possibilities for non-linear mixer action in this tube, the result of which is not included in the input voltage value given by the calibration curve. It is recommended that the first i-f tube is operated as linearly as possible.

The i-f voltmeter must be free from feedback so that, if the amplifier is calibrated from a zero-impedance source, the calibration remains true even when the input impedance is increased to a high value during the practical measurements. For this reason extensive decoupling and shielding may be required.

MEASUREMENT RESULTS

Several sets of conversion diagrams were obtained under various operational conditions; each set comprising four to six diagrams. As an example two diagrams out of a series of six for different values of grid bias battery voltage $\bar{E}_{cc}$ are shown in Fig. 4. These diagrams are obtained with a resistor $R_0 = 5000$ ohms in the cathode lead and therefore describe conditions with mixed automatic and fixed bias voltage. The B-wave voltage is constant and equal to 1.5 volts r.m.s, and the plate bias voltage is 50 volts. The measurement frequency is $A/2\pi = 7$ mc.p.s.

As is generally the case, the area of a conversion diagram indicates the amount of i-f output power, while the distribution of area describes the matching conditions. A comparison of the two diagrams in Fig. 4 shows that a bias battery voltage of -3 volts gives small i-f output power; the maximum power being obtained when the load impedance is approximately 60,000 ohms. A bias battery voltage of $+3$ volts gives considerably more i-f output power, the maximum now being obtained when the load impedance is approximately 20,000 ohms. These conditions are more clearly brought out by the dotted

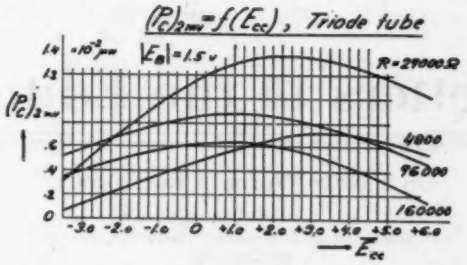


FIG. 5. Curves showing the i-f output power P_C as function of the battery bias value $\bar{E}_{cc}$ with a cathode resistor $R_0 = 5000$ ohms and an oscillator voltage $E_B = 1.5$v r.m.s.

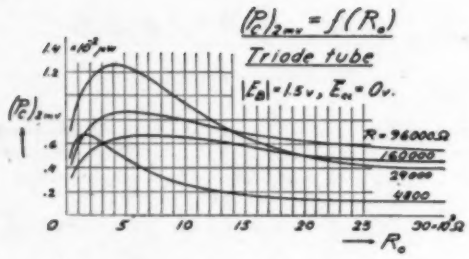


FIG. 6. Curves showing the i-f output power P_C as function of the cathode resistor value R_0 with zero battery bias and an oscillator voltage $E_B = 1.5$v r.m.s.

power curves for signal input voltages of 2.0 mv and 1.6 mv r.m.s. value.

A better study of the biasing conditions can be made if separate curves are plotted from suitable sets of conversion diagrams. Figure 5 shows a family of curves obtained from a series of diagrams including the two in Fig. 4 with a cathode resistor value $R_0 = 5000$ ohms. It is seen that for maximum i-f output power, the converter should be operated with a positive battery bias of the order of two volts and with a load impedance of the order of 30,000 ohms. The flatness of the curve arises from the fact that a change in battery bias voltage is automatically counteracted by the resulting change in the bias voltage developed across the resistor R_0. The family of curves in Fig. 6 are obtained from a set of conversion diagrams for zero-bias battery voltage and the same load impedances as in the previous case. Here the cathode resistor R_0 is varied, and it is seen that approximately 5000 ohms gives maximum i-f output power when the load impedance has the optimum value of approximately 30,000 ohms. Higher and lower values of load impedance give less output, but also less dependence upon the actual value of the biasing resistor.

It is understood that the above diagrams and curves do not completely describe the operation of a triode tube mixer but merely serve to indicate how the operation may be studied by means of this method. The treatment may be extended to the determination of variational conversion coefficients from the diagrams in Fig. 4 and calculations with the aid of formulas similar to the ones derived for crystal and diode tube mixers. Equivalent circuits representing these formulas can also be drawn.[1,2]

CONCLUSIONS

The mathematical treatment of triode mixers becomes very involved when the frequency is high and the operation extremely non-linear ("Class B" operation). Frequency conversion diagrams provide a shortcut around extensive calculations and changes the nature of experimental investigation to become a standardized procedure. Although in the conversion diagrams shown in the text the information refers to the non-linear element plus load, it is sometimes desirable that part of the practical input or output circuit is also included. If this is done the diagrams can be made to describe the total characteristics of a practical frequency converter, but the data are then of less general importance when the non-linear element is used in other types of circuits. Still there is a number of additional design data and requirements to be taken into account, for example requirements set by the signal-to-noise ratio. Much experience is required before all the additional design data can be dealt with successfully. The conversion diagram does not solve all converter design problems, but provides the designer with a tool for recording some of the fundamental relations governing the behavior of the practical frequency conversion circuit.

ACKNOWLEDGMENTS

The author wishes to thank Professor E. L. Chaffee, Director of Cruft Laboratory, for use of his idea of the frequency conversion diagram.

Letters to the Editor

On a Novel Form of Refrigerator

W. D. HAYES
Lockheed Aircraft Corporation, Burbank, California
July 30, 1945

IN a recent paper[1] in this journal, Mr. J. R. Roebuck proposed a system of refrigeration and calculated the coefficient of performance or "refrigeration efficiency" for this system. This quantity he compared with that for a reversed Carnot cycle, and arrived at the conclusion that the refrigeration cycle he proposed holds a very material advantage over the reversed Carnot cycle.

This letter reveals a fundamental fallacy in the reasoning behind the comparison of the two refrigeration cycles in this paper. Figure 1 is presented to exhibit the cycles compared on a temperature-entropy plot. The points α, β, γ, and δ outlining Mr. Roebuck's proposed cycle are the same as in Fig. 4 of the original paper. The rectangle completed by the dotted line is the reversed Carnot cycle.

Refrigeration is accomplished only when a quantity of heat is transferred from a low temperature heat reservoir

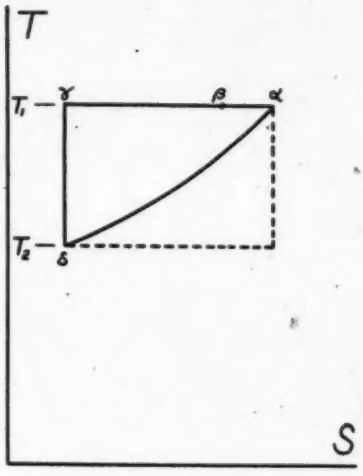


FIG. 1.

to a reservoir at higher temperature. The reversed Carnot cycle considered transfers a quantity of heat $T_2\Delta S$ from a reservoir at temperature T_2 to a reservoir at temperature T_1. In the proposed refrigerator cycle, no heat can be taken from a reservoir at temperature T_2, and the quantity of heat absorbed by the system cannot be taken from any reservoir of temperature less than T_1. The refrigeration system for which the efficiency was calculated therefore operates between temperatures T_1 and T_2, and accomplishes no refrigeration. The reversed Carnot cycle that should be compared with the proposed cycle would be represented by a straight line (α to γ to α) in Fig. 1 and would have an *infinite* coefficient of performance.

Mr. Roebuck's system as proposed would be capable of refrigeration in a counterflow heat exchanger, where the temperature of the refrigerated medium is not uniform. It would be capable of refrigeration from a heat reservoir of uniform temperature were the auxiliary compression carried out isentropically rather than isothermally, making the line $\alpha\beta$ in Fig. 1 vertical.

[1] J. R. Roebuck, "A Novel Form of Refrigerator," J. App. Phys. 16, 285 (1945).

Wave Equations for Finite Elastic Strains

HUBERT M. JAMES
Department of Physics, Purdue University, Lafayette, Indiana
AND
EUGENE GUTH
Department of Physics, Notre Dame University, Notre Dame, Indiana
July 23, 1945

IN a recent note in this Journal, Dr. A. Nadai[1] comments on a paper by B. A. Mrowca, S. L. Dart, and E. Guth[2] and on a note by H. M. James and E. Guth.[3] Dr. Nadai criticizes the wave equation given by James and Guth for longitudinal vibrations in a slender uniform rod obeying Hooke's law, as being valid for infinitesimal strains only. This equation is

$$\rho\frac{\partial^2 u}{\partial t^2}=f\frac{\partial^3 u}{\partial^2 x\partial t}+E\frac{\partial^2 u}{\partial x^2}, \tag{1}$$

where x measures distance along the bar to a point in the unstressed bar, u is the displacement of this point when the bar is in vibration, ρ is the volume density of the bar, f is a damping constant, E is Young's modulus, and t is time. Instead of this, Dr. Nadai proposes a non-linear wave equation which becomes, for negligible damping,

$$\left(1+\frac{\partial u}{\partial x}\right)^{-3}\frac{\partial^2 u}{\partial x^2}=\frac{\rho}{E}\frac{\partial^2 u}{\partial t^2}. \tag{2}$$

This he considers to be valid for finite strains. It appears to us, however, that there is a fundamental error in the derivation of Dr. Nadai's equation, and that it is our Eq. (1), and not his Eq. (2), which is valid for finite strains. In order to clarify this point it is necessary to carry through the first stages in the derivation of these equations in terms of the alternative spatial variables x and $x'=x+u$, which gives the position of a point in the stretched bar. (For a given element of the bar x is thus independent of time, while x' is variable.)

We consider the infinitesimal section of the bar which in its unstretched state extends from x to $x+dx$, in its stretched state from x' to $x'+dx'$. Its length is

$$dx'=(1+\partial u/\partial x)dx=(1+\epsilon)dx, \tag{3}$$

ϵ being the local elongation, in general a function of x or x', as well as of t. Assuming that ρ does not change on stretching, one has for the cross section of the distorted bar

$$s=s/(1+\epsilon), \tag{4}$$

s being the small constant cross section of the unstretched bar. The mass of the infinitesimal section of the rod under consideration is

$$dm=\rho s dx=\rho s(x',t)dx'. \tag{5}$$

Let $\sigma(x,t)$ be the nominal stress (the force per unit original cross section) and $\bar{\sigma}(x',t)$ the true stress (the force per unit actual cross section). Then as alternative expressions for the net force on this section of the bar we have

$$dF=s\sigma(x+dx)-s\sigma(x)\rightarrow s\frac{\partial\sigma}{\partial x}dx, \tag{6}$$

or

$$dF=s(x'+dx')\bar{\sigma}(x'+dx')-s(x')\bar{\sigma}(x')\rightarrow\frac{\partial(s\bar{\sigma})}{\partial x'}dx'. \tag{7}$$

The acceleration of this section of the bar is $a=(\partial^2 u)/(\partial t^2)$. Writing $dF=dm\cdot a$ one has on combining Eqs. (5) and (6)

$$\frac{\partial\sigma}{\partial x}=\rho\frac{\partial^2 u}{\partial t^2}, \tag{8}$$

and on combining Eqs. (5) and (7)

$$\frac{1}{s}\frac{\partial(s\bar{\sigma})}{\partial x'}=\rho\frac{\partial^2 u}{\partial t^2}. \tag{9}$$

However, Dr. Nadai starts his derivation from the relation

$$\frac{\partial\bar{\sigma}}{\partial x'}=\rho\frac{\partial^2 u}{\partial t^2}. \tag{10}$$

Thus Dr. Nadai neglects a generally non-zero term, $(\bar{\sigma}/s)(\partial s/\partial x')$, on the left. This is of the second order of smallness for infinitesimal strains, but its neglection does limit the validity of Dr. Nadai's equations to the case of infinitesimal strains. Our derivation starts from Eq. (8), and its validity is not thus limited.

One other difference exists between our treatment and that of Dr. Nadai. As "Hooke's law" we write

$$\sigma=E\epsilon=E\frac{\partial u}{\partial x}. \tag{11}$$

Combination of this with Eq. (8) leads to Eq. (1) above, with the damping term neglected. Dr. Nadai instead introduces the natural strain

$$\bar{\epsilon}=\ln(1+\epsilon), \tag{12}$$

and takes as "Hooke's law"

$$\sigma = E\epsilon. \tag{13}$$

Neither of these equations represents exactly the behavior of *real* rubber at *large* strains, which is more accurately given[4] by

$$\sigma = \frac{E}{3}[1+\epsilon-1/(1+\epsilon)^2]. \tag{14}$$

As descriptions of real rubber both Eqs. (11) and (13) thus neglect terms of the order of ϵ^3. Choice between these approximations in treating finite but *not too large elongations* does not seem to us to be a matter of correctness, but of naturalness and convenience. Both these considerations strongly indicate the use of Eq. (11), which leads to soluble differential equations, whereas use of Eq. (13) leads to non-linear equations of complicated and forbidding appearance.

Neither our form of Hooke's law nor Dr. Nadai's gives a very satisfactory approximation for real rubber at elongations of the order of 100 percent. On combining Eqs. (8) and (14) one obtains a non-linear wave equation which should be satisfactory for this and even greater elongations:

$$\rho\frac{\partial^2 u}{\partial t^2} = \frac{E}{3}\left[1+\frac{2}{(1+\partial u/\partial x)^3}\right]\frac{\partial^2 u}{\partial x^2}. \tag{15}$$

This equation is formally similar to the equation for sound waves of finite amplitude, and is correspondingly hard to handle.

Dr. Nadai states that in reference 2 the wave equation (1) of the present note was compared with experimental data obtained at extensions of 100 percent. It is a matter of fact, however, that in reference 2 Eq. (1) was used only for smaller extensions. For larger extensions a wave equation exactly equivalent to Eq. (15) of the present note was given in reference 2 [cf. Eqs. (12)–(14) of reference 2].

[1] A. Nadai, J. App. Phys. 16, 313 (1945).
[2] B. A. Mrowca, S. L. Dart, and E. Guth, J. App. Phys. 16, 8 (1945).
[3] H. M. James and E. Guth, Phys. Rev. 66, 33 (1944).
[4] H. M. James and E. Guth, J. Chem. Phys. 11, 455 (1943).

Recent Application of Physics

Prepared by CLARK GOODMAN, Associate Editor

Material to be included in this section should be submitted to Dr. Clark Goodman, Massachusetts Institute of Technology, Cambridge 39, Massachusetts.

Ultrasonics in High Polymer Research[1]

It has been demonstrated that ultrasonic vibrations can produce physical and chemical changes in high molecular weight polymers. Solutions of nitrocellulose, polyvinylacetate, and polystyrene undergo irreversible decreases in viscosity when subjected to about 10 watts per cm² at 300 kc for only a few minutes. Apparently the larger molecules (molecular weight $>10^5$) are preferentially broken by the frictional force developed between themselves and the solvent, resulting in a more homogeneous product of lower molecular weight and hence lower viscosity. The strengths of covalent C−C and C−O bonds, based on infra-red spectroscopic measurements, are approximately equal to the forces to be expected from such ultrasonic vibrations. Using about the same intensity but somewhat lower frequencies (20 to 100 kc), the forces are insufficient to rupture the intramolecular bonds but can open the intermolecular van der Waals bonds which give high apparent viscosity to thixotropic solutions of certain polymers. The effect with these materials is similar to that of mechanical agitation, for after the sonic vibration stops, the irregular Brownian motion reestablishes the random molecular network in the solution and the original viscosity returns.

[1] Brief review by H. Mark, J. Acous. Soc. Am. 16, 183–187 (1945).